



TIETO- JA SÄHKÖTEKNIIKAN TIEDEKUNTA

Markus Heino

**Vedenkulutusdatan visualisointi ja hyödyntäminen
vesihuoltolaitoksille suunnatussa Datapalvelussa**

Diplomityö
Tietotekniikan tutkinto-ohjelma
Helmikuu 2020

Heino M. (2020) Vedenkulutusdatan visualisointi ja hyödyntäminen vesihuoltolaitoksille suunnatussa Datapalvelussa. Oulun yliopisto, tietotekniikan tutkinto-ohjelma. Diplomityö, 61 s.

TIIVISTELMÄ

Suomen vesihuolto on siirtymävaiheessa kohti digitalisaatiota. Moni vesihuoltolaitos on aloittanut etäluettavien mittareiden kokeilun ja toisilla etäluettaviin mittareihin siirtyminen on suoritettu täysin tai osittain. IoT-mittarit lähettävät valtavat määrät dataa, jonka tallentamiseen, hallitsemiseen ja analysointiin kaivataan työkaluja.

Jotta siirtymävaihe onnistuisi mutkattomasti, vesihuollon parissa kartoitetaan verkkoteknologioita, eri mittarinvalmistajia sekä tehdään lukuisia tutkimuksia, kuinka digitalisointi tullaan toteuttamaan. Tutkimuksissa selvitetään, mihin ja kuinka etäluettavista mittareista saatu data voidaan tallentaa ja muuttaa informatiiviseksi. Vesihuolto on tilanteessa, jossa se tarvitsee järjestelmän, joka pystyy tukemaan datan tallentamisen, hallitsemisen ja analysoimisen. Ongelmaan ratkaisuna on lähdetty rakentamaan Vesitieto Oy:n toimesta Datapalvelua.

Työ tutustuttaa lukijan, miten Datapalvelu on rakennettu ja mitä mahdollisia ratkaisuja Datapalvelu tarjoaa mainittuihin ongelmiin. Tämän työn parissa keskitytään erityisesti IoT-mittarista saadun datan hyödyntämiseen, analysoimiseen sekä visualisoimiseen Datapalvelua hyödyntäen. Lukijalle työ tarjoaa kattavan katsauksen massadatasta, tutkimusmenetelmistä, IoT:sta, datan visualisoinnista ja visualisointiin olevista työkaluista sekä yleiskuvan vesihuollon tilanteesta ja tulevaisuudesta.

Avainsanat: Esineiden internet, IoT, massadata, big data, älykkäät mittarit, ohjelmisto

Heino M. (2020) Visualization and utilization of water consumption data in a Dataservice for water supply plants. University of Oulu, Degree Program in Computer Science and Engineering. Master's Thesis, 61 p.

ABSTRACT

The Finnish waterworks are in the transition phase to digitalization. Many waterworks have begun experimenting with smart water meters when others are already changing their meter towards smart water meters. IoT-meters send tremendous amounts of data that require tools to store, manage and analyze them.

To make digitalization a success, water management is mapping network technologies, different meter manufacturers, and conducting numerous studies on how to make the transition smooth toward digitalization. The studies will focus on where and how remotely collected data could be stored and transformed into information. Water supply is in a situation where it needs a system that can support data storage, management, visualization and analysis. To find a solution to the problem Vesitieto Oy has built Dataservice.

The work introduces the reader to how the Dataservice is constructed and what possible solutions it offers to the problems mentioned. This work focuses particularly on utilizing, analyzing and visualizing data gathered from smart water meters using Dataservice. The work provides the reader with a comprehensive overview of mass data, research methods, IoT, data visualization and visualization tools, and an overview of the state and future of water supply.

Keywords: IoT, Big data, Smart water meter, software

SISÄLLYSLUETTELO

TIIVISTELMÄ

ABSTRACT

SISÄLLYSLUETTELO

ALKULAUSE

LYHENTEIDEN JA MERKKIEN SELITYKSET

1.	JOHDANTO	8
1.1.	Tutkimuksen tausta	8
1.2.	Tutkimuksen tavoite	8
1.3.	Aiheen rajaus	9
1.4.	Vesitieto Oy	9
2.	TEORIA JA MENETELMÄT	11
2.1.	Internet of Things	11
2.2.	Teollinen internet	11
2.2.1.	LoRa	12
2.3.	Kontekstuaalinen suunnittelu	12
2.3.1.	Asiakkaaseen liittyvä tiedustelu	12
2.4.	Kvalitatiivinen data ja analyysi	13
2.4.1.	Henkilökohtainen haastattelu	15
2.4.2.	Kohderyhmähaastattelu	16
2.4.3.	Havainnointi ja tarkkailu	17
2.4.4.	Tapaustutkimus	17
2.5.	Kvantitatiivinen data ja analyysi	18
2.5.1.	Mittaus ja virheet	18
2.6.	Visualisoituminen käsitteenä	19
2.7.	Visuaalinen tiedustelu	20
2.8.	Selkeä visualisointi	20
2.9.	Visualisoinnin tavoite	21
2.10.	Työkalut visualisointiin	22
2.10.1.	Pylväskaavio	23
2.10.2.	Viivakaavio	27
2.10.3.	Pylväs- Viivakaavio	28
2.10.4.	Piirakkakaavio	29
2.10.5.	Tutkakaavio	29
2.10.6.	Datakartat	30
2.10.7.	Bullet-kaavio	32
2.10.8.	Vesiputous-kaavio	32
2.10.9.	Laatikkokaavio	34
2.11.	Kirjastot ohjelmointiin	35
2.11.1.	Chart.js	35
2.11.2.	D3.js	35
2.11.3.	Dygraphs	35
2.12.	Mitattavat laitteet	36
2.13.	Vastaanotettu data	36
2.13.1.	Raakadata ja tietämys	37
2.13.2.	Rakenne	37

2.14.	Massadata	38
3.	VESIHUOLTOLAITOKSET	40
3.1.	Vesihuoltolaitokset vuonna 2019	40
3.2.	Vesihuollon säännökset	40
3.3.	Vesihuollon tarpeet tulevaisuudessa	41
4.	DATAPALVELU TOTEUTUS	43
4.1.	Yhteistyö	43
4.2.	Ensisijaiset tarpeet	44
4.2.1.	Toteutus prioriteetit	45
4.3.	Määrittely	45
4.3.1.	Tietokanta	46
4.3.2.	API	46
4.3.3.	Verkko	47
4.3.4.	Kerätty data ja sen hyödyntäminen	47
4.4.	Käyttöliittymä	47
4.4.1.	Visualisointi ja työkalut	49
4.4.2.	Kulutuksen seuranta	50
4.4.3.	Hälytykset	52
4.4.4.	Alueiden määrittely piirtämällä	53
4.5.	Tulevaisuus	54
5.	TULOSTEN TARKASTELU	55
6.	YHTEENVETO	57
7.	LÄHTEET	58

ALKULAUSE

Erityiset kiitokset työn ohjauksesta toteutus- ja kirjoitusprosessin aikana kuuluvat Pekka Siirtolalle ja Eija Ferreiralle. Kiitokset myös Oulun Veden Tero Kilpeläiselle sekä Sara Alanärälle. Työn aiheesta ja työympäristön tarjoamisesta suuri kiitos Suomen Vesitieto Oy:lle ja erityisesti Markku Ojalalle. Lisäksi kiitos yksilöille, jotka ovat auttaneet työn parissa sekä tarjonneet ohjausta, tukea ja palautetta projektin aikana. Suuri kiitos myös työparilleni Jussi Jokiolle, joka työstää Datapalvelun palvelinpuolta (Back-End).

Oulu, Helmikuu 6. 2020

Markus Heino

LYHENTEIDEN JA MERKKIEN SELITYKSET

AI artificial intelligence, tekoäly

API application programming interface, sovellusohjelmointirajapinta

IoT internet of things, esineiden internet

IIoT industrial internet of things, teollinen internet.

kpl kappale

ML machine learning, koneoppiminen

LoRa long range, pitkän kantaman

UI user interface, käyttöliittymä

kk kuukausi

js javascript

1. JOHDANTO

1.1. Tutkimuksen tausta

Vesilaitokset ovat osoittaneet intoa siirtyä enemmän kohti digitalisoitumista. Tämä on tullut esille muun muassa AIF Water Ecosystem hankkeessa EU-Horizon: Datan hallinta & Tiedon hyödyntäminen. Pirkko Taskinen on vetäjänä kyseisessä hankkeessa ja on järjestänyt työpajoja, joissa on ollut osallisena muun muassa vesilaitosten edustajia, yliopiston tutkijoita, mittalaitteiden toimittajia ja alan yrityksiä sekä eri vesihuollon toimijoita. Hankkeessa on pyritty hakemaan rahoitusta vesilaitosten digitalisointiin ja perehdytty asioihin, joihin on kiinnitettävä eniten huomiota ja jotka ovat prioriteetiltaan tärkeimpiä asioita vesilaitoksille, kun digitalisaatio saadaan kunnolla aloitettua.[1]

Vesitieto Oy on yksi työpajoihin osallistuvista yrityksistä, joka on tehnyt asiakas- ja laskutusjärjestelmän vesilaitoksille, kuten Limingan vesihuollolle, Oulun Vedelle sekä muutamalle muulle vesihuoltolaitokselle. Vesilaitokset ovat erittäin kiinnostuneita siirtymään vaiheittain kohti etäluettavia mittareita. Etäluettavista mittareista saadun datan tallentamiseen ja hyödyntämiseen tarvitaan uusia työkaluja erityisesti vesihuollon parissa. Vesilaitokset eivät ole halukkaita itse kehittämään kyseisiä työkaluja, vaan haluavat ostaa palvelun tarjoajia, jotka pystyvät toimittamaan ratkaisuja. Datapalvelun rakentaminen ja datan visuaalinen esitys on tuntunut luontevalta seuraavalta askeleelta Vesitiedolle, ja näin ollen tutkimustyötä on aloitettu tekemään, jotta kehitystyö saadaan aloitettua.[2]

1.2. Tutkimuksen tavoite

Tutkimuksen tavoitteena on löytää ratkaisuja sekä menetelmiä Datapalvelun rakentamiseen ja sen maksimaaliseen hyödyntämiseen. Menetelmien tulee voida kerätä, hallita ja hyödyntää massadataa. Erityisesti tärkeää on, että jokainen vesihuollon parissa toimiva taho pystyy hyödyntämään Datapalvelua tiedon tallentamiseen, keskittämiseen, keräämiseen, visualisoimiseen tai osaan näistä.[2]

Datapalvelun suunnittelu ja toteutus kohdistuu ensisijaisesti vesihuoltolaitoksille ja etäluettaviin mittareihin. Haastattelujen, kokouksien ja Vesitieto Oy:n tarjoaman tiedon ja tietämyksen kautta keskitymme niihin osa-alueen ongelmiin, jotka ovat tällä hetkellä sekä lähitulevaisuudessa prioriteetiltaan korkeimmalla vesihuoltolaitoksille.

Tiivis yhteistyö eri sidosryhmien kanssa koko projektin ajan ja myös sen jälkeen nopeuttaa projektin suunnittelua sekä toteutusta. Tiiviin yhteistyön ansiosta saadaan välitöntä palautetta sidosryhmiltä, jotka heijastuvat suunnitteluun ja toteutukseen. Kontekstuaalisen suunnittelun avulla sidosryhmät ovat lähellä projektia alusta lähtien. Kontekstuaalista suunnittelua hyödyntäen Vesitieto pääsee lopputuotteeseen, jossa asiakkaiden tarpeisiin pystytään tarjoamaan palveluita ensimmäisistä toimivista versioista lähtien. Suunnittelussa on otettu huomioon myös mahdollinen laajentuminen muille toimialueille tulevaisuudessa.

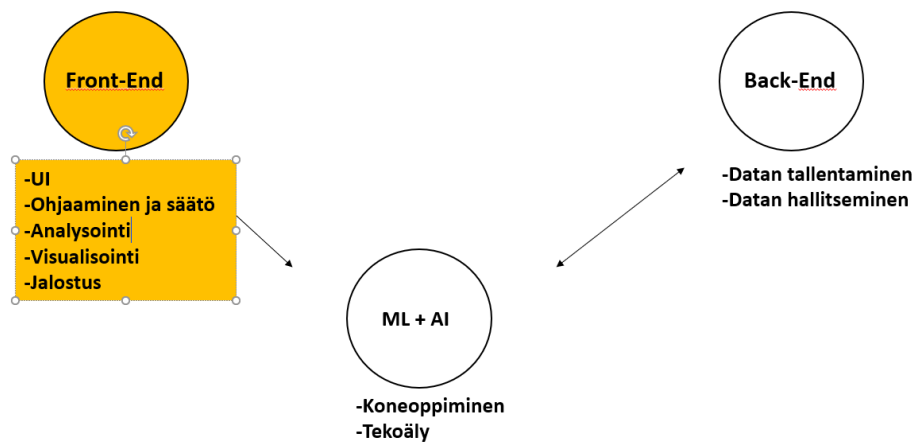
Datapalvelun on tarkoitus tarjota tuote massadatan keräykseen, sen tallentamiseen ja analysoimiseen. Datapalvelun on tarkoitus hyödyntää myös kolmansilta osapuolilta saatua tietoa tarkempaa ja tehokkaampaa analyysia varten. Tarjottujen palveluiden avulla Datapalvelu pystyy kohdistamaan tuotettua informaatiota ja tietoa asiakkaille entistä paremmin ja laajemmin. Rajapintoja hyödyntäen on tarkoitus kerätä

esimerkiksi asiakas- ja laskutustietoja Vesitiedon palvelusta. Tätä tietoa hyödyntämällä pyritään tarjoamaan asiakkaille mahdollisimman laajaa kokonaiskuvaa omasta toiminnasta, sen seuraamisesta sekä parantamisesta.

1.3. Aiheen rajausta

Datapalvelua ja siihen kuuluvaa kokonaisuutta on pilkottu pienempiin osiin. Tämän työn osalta projekti keskittyy yhteen kolmesta osasta. Kuva 1 on leikkaus isommasta kokonaisuudesta, jonka on suunnitellut Vesitieto Oy. Kuvasta 1 nähdään Datapalvelun keskittyvän datan hallitsemiseen, tallentamiseen, visualisoimiseen sekä jalostamiseen. Erityisesti tämä katsaus keskittyy Front-End puoleen, kerätyn datan hyödyntämiseen ja jalostamiseen (Kuva 1, keltaisella korostetut alueet). Työ keskittyy UI-toteutukseen, datan analysointiin, sen visualisointiin ja jalostukseen. Lyhyesti sanottua työ keskittyy siihen, kuinka muutamme massadatan informaatioksi, siitä tiedoksi ja lopuksi tietämykseksi hyödyntäen niin kvalitatiivista kuin kvantitatiivista dataa. Jotta tämä tavoite saavutettaisiin, on tämä työ jaettu seuraaviin osatavoitteisiin.

- UI kehitys ja suunnittelu
- Datan analysointi
- Datan visualisointi
- Datan jalostus



Kuva 1. Tutkimusprojektin rajausta.[2]

1.4. Vesitieto Oy

Suomen Vesitieto Oy on oululainen yritys, joka on perustettu vuonna 2017. Vesitieto kehittää ja rakentaa älykkäitä ratkaisuja vesihuollon digitalisointiin. Kehityksessä

käytetään hyväksi IoT-mittauksia ja analytiikkaa. Järjestelmät on toteutettu yhteistyössä vesihuoltolaitosten kanssa ja ratkaisut ovat pääosin web-pohjaisia.

Ensimmäistä tuotetta vesihuollon piiriin on alettu kehittämään jo vuonna 2015. Asiakkaille on toimitettu ensimmäiset toimivat ratkaisut vuonna 2018, jolloin voidaan sanoa myös liiketoiminnan alkaneen. Vesitiedon asiakkaisiin kuuluu Suomen suurimpia vesilaitoksia kuten Oulun Vesi, ja toimituksia on vireillä useisiin eri kaupunkeihin tälläkin hetkellä (2019). Liikevaihdon on arveltu yltävän vuonna 2019 miljoonan euron luokkaan, yrityksen työllistäessä kymmenen henkilöä.

Liiketoiminnan ydin keskittyy pääosin vesihuollon pariin. Vesitiedon suunnitelmissa on laajentaminen kaukolämmön, jätehuollon ja kaasuyrityksien puolelle. Tulevaisuuden suunnitelmissa on mainittu myös muu energiateollisuus kuten sähkö, sekä levittäytyminen Suomen rajojen ulkopuolelle.[2]

Datapalvelu oli seuraava askel Vesitiedolle ja sitä tullaan tarjoamaan nykyisille sekä uusille asiakkaille sen valmistuttua. Tarkempaa aikataulua valmistumiselle ei ole määritelty, mutta ensimmäisiä demoversioita Datapalvelusta on näytetty nykyisille asiakkaille marraskuun alusta 2019.

2. TEORIA JA MENETELMÄT

Yhdistämällä useita teknologioita on mahdollista kerätä dataa ympäristöön sulautetuista laitteista. Kerättyä dataa voidaan hyödyntää, kun tehdään arvioita, ennustuksia, ja kun halutaan reagoida erilaisiin muutoksiin tehokkaasti ja nopeasti. Mikäli kerättyä dataa on tarpeeksi, voidaan hyödyntää tekoälyä, jolloin järjestelmät ja laitteet voivat tehdä itsenäisiä päätöksiä ilman, että tarvitaan ihmistä. Tekoälyn tuottaman tiedon ja ratkaisujen avulla on mahdollista reagoida muuttuviin olosuhteisiin nopeammin. Tämän kaltaista ympäristöä ja vuorovaikutusta kutsutaan älykkääksi ympäristöksi. Älykkäästä ympäristöstä puhuttaessa IoT (Internet of Things, esineiden internet) on yksi merkittävä tekijä.

2.1. Internet of Things

Internet of Things -termiä ehdotti vuonna 1999 Kevin Ashton, ja häntä pidetään käsitteen keksijänä [3, 4]. Puhuttaessa IoT:sta tarkoitetaan laitteita, jotka kykenevät toimimaan sekä lähettämään dataa. Oleellista myös lähetettävän datan kannalta on, että dataa voidaan myös hyödyntää jollakin tavalla. Tämä on mahdollista, kun on olemassa useita eri teknologioita ja niitä voidaan yhdistellä. Näin voidaan luoda iso toimiva kokonaisuus. Laitteet ja teknologiat, jotka on pystytty yhdistämään, muodostavat suuren älykkään ympäristön, jota hallitaan Internet-yhteyttä hyödyntäen. Laitteet lähettävät dataa, joka kerätään talteen ja tallennetaan järjestelmään. Tallennettua dataa voidaan hyödyntää sekä jalostaa monenlaisiin tarkoituksiin riippuen datasta, sen laadusta ja ongelmista, joihin halutaan ratkaisu.[5]

2.2. Teollinen internet

Teollinen internet, joka tunnetaan globaalisti nimellä Industrial Internet of Things (IIoT), tarkoitetaan yritysten digitalisoitumista. Yrityksen digitalisoitumisen myötä pyritään tehostamaan ja kehittämään omia palveluita eteenpäin. Kerättyä dataa tehokkaasti ja runsaasti yrityksen käytössä olevista laitteista, koneista, antureista ja sensoreista, voidaan päätellä ja tehdä arvioita tarkemmin etukäteen ja tulevaisuuteen. Analysoimalla kerätty massadata voidaan parantaa yrityksen tehokkuutta monilla, ellei jopa kaikilla liiketoiminnan kriittisillä osa-alueilla. Lisäksi voidaan optimoida yrityksen toimintaa, jolloin tarkoituksena on tuottavuuden paraneminen, talouden kasvun nopeutuminen, tulevaisuuden ennustamisen tarkentuminen ja lopputuloksena markkina-aseman vahvistuminen.[6]

IIoT:n on mahdollistanut teknologioiden kehityksen monilla osa-alueilla muun muassa, pilvilaskenta (cloud computing) [7], mobiiliteknologia, kyberturvallisuus, massadata (Big Data), IoT, koneoppiminen (machine learning lyhennettynä ML) ja tekoäly (Artificial Intelligence lyhennettynä AI eli tekoäly) [8]. IIoT vaatii kaikkia edellä mainittuja teknologioita toimiakseen ja tuottamaan hyödyllistä analytiikkaa yrityksille. Lyhyesti sanottuna teollinen internet on fyysisten laitteiden liittäminen toisiinsa verkossa, jossa data voidaan siirtää eteenpäin turvallisesti eri protokollia ja verkkoratkaisuja hyödyntäen, ajatuksena parantaa tuottavuutta ja tehokkuutta.[9, 10]

2.2.1. LoRa

LoRa tulee englannin kielisistä sanoista “Long Range” ja sillä tarkoitetaan pitkän kantaman ja pienen tehon kuluttavaa laitetta, joka käyttää hyväkseen langatonta radioyhteyttä. Ensimmäiset LoRa laitteet on kehitetty vuonna 2012 Cycleo of Grenoblen toimesta. LoRa teknologiaa hyödyntäen voidaan data muun muassa laitteista ja sensoreista lähettää eteenpäin.[11]

LoRa käyttää radioverkon taajuutta 433-, 868-, 915-megahertsiä (MHz) Euroopan alueella. Lähetys kantama voi olla jopa yli 10 kilometriä maastosta riippuen. Vuonna 2018 LoRa laitteiden ilmoitettiin kuluttavan entistä vähemmän tehoa. Samana vuonna lähetyksen kantama on lisääntynyt ja laitteen koko pienentynyt verrattuna aikaisempiin laitteisiin.[11]

LoRa teknologiaa käytetään hyväksi lämpöenergian puolella ainakin Oulun alueella. Laitteita, jotka käyttävät LoRa teknologiaa hyödyksi esimerkiksi Oulun alueella, ovat etäluettavat lämpömittarit. Etäluettavia lämpömittareita, jotka hyödyntävät LoRa teknologiaa on Oulun alueella noin 30000 kappaletta.

2.3. Kontekstuaalinen suunnittelu

Kontekstuaalinen lähestymistapa tarkoittaa suunnittelua asiakkaan näkökulmasta ja asiakkaan toimintatavasta.

- Tavoitteena on kerätä tietoa siitä, mitä asiakas tarvitsee
- Tiedustella markkinoita ja mahdollisia käyttäjäryhmiä
- Tyydyttää asiakkaan tarpeet uudistuneella suunnittelulla
- Suunnitella systeemi, joka tyydyttää käyttäjäryhmän tarpeet
- Testata ympäristöä nopeasti käyttäjäryhmien kanssa

Kontekstuaalisen suunnittelun kautta kehittäjä saa yksityiskohtaisen kuvan siitä, mitä asiakas tarvitsee. Toisaalta kehittäjä saa ymmärryksen vaaditusta teknologiasta, jota suunnittelu sekä toteutus vaatii. Lähtökohta on kuitenkin asiakkaan kanssa tiivis yhteistyö, jota hyödyntäen pystytään mukautumaan asiakkaan tarpeisiin nopeammin ja joustavammin.[12]

2.3.1. Asiakkaaseen liittyvä tiedustelu

Kehittäjän työ on kerätä informaatiota asiakkaan tarpeista ja haluista. Erityisesti on tärkeää tiedustella, kuinka asiakas hoitaa oman työnsä ja minkälaisia työkaluja käyttäen. Kaikkien tarpeiden kuvaaminen sanoin on kuitenkin usein vaikeaa. Tämä johtuu osittain työympäristöstä ja totutuista tavoista, jota työympäristössä käytetään. Toisaalta työntekijät eivät aina ymmärrä minkälainen vaikutus heidän oman työnsä tekemisellä on suuremmassa mittakaavassa.[12]

Tiedon keräämiseksi on haastateltava useaa eri asiakasta saman ongelman piirissä. Näin pystytään keräämään suuri määrä informaatiota tarpeista ja haluista, jotka voidaan yhdistää. Näin voidaan muodostaa kokonaiskuva potentiaalisten asiakkaiden markkina alueesta, heidän keräämästään informaatiosta ja liiketoiminnasta. Kerätyn tiedon avulla suurempien ja pienempien yksittäisten asiakkaiden tarpeet eivät jää huomioimatta. On mahdollista myös tunnistaa tiettyjä tapoja, joita kyseisen

liiketoiminnan parissa käytetään usein. Asiakkaan liiketoimintaa pyörittävät ratkaisut ovat usein jo toimivat ja tuloksellisia. Tarkoituksena tiedustelun avulla on kuitenkin tuottaa järjestelmä, joka tehostaa asiakkaan liiketoimintaa useilla eri osa-alueilla. Tiedustelun onkin syytä keskittyä asiakkaan järjestelmiin, jotka eivät toimi tehokkaasti tai järjestelmiin, joiden tehokkuutta voidaan parantaa huomattavasti uudella innovatiivisella järjestelmällä.[12]

Usein tuote tai järjestelmä tehdään yhdelle asiakkaalle. Haaste muodostuu, kun tuotteen tilaavan asiakkaan tarpeet nousevat esiin, toisin kuin muiden potentiaalisten asiakkaiden. Tällöin on mahdollista, että suunnittelu ei toteudu kunnolla ja järjestelmä ei skaalaudu uusille asiakkaille samalla tapaa ja tuota haluttua tehokkuutta. Riski tapahtuu lähes vääjäämättä, mutta oleellista suunnittelussa on ottaa huomioon myös tulevien asiakkaiden tarpeet ja halut ainakin osittain. Järjestelmään suunniteltaessa on otettava huomioon tulevaisuus. Tulevaisuudessa on pystyttävä tarjoamaan yhä enemmän ja laajempia palveluita niin uusille kuin vanhoille asiakkaille.[12]

2.4. Kvalitatiivinen data ja analyysi

Datalla on monta eri muotoa, se voi olla numeerista, tietämystä, tutkimusta, digitaalista tai jotain muuta. Käytännössä dataa on kaikki mahdollinen tieto. Tämän takia datalle on myös monia eri analysointimenetelmiä. Yhdistelemällä eri tietolähteitä saadaan uutta tietoa ja tarvitaan taas uusia menetelmiä analysointiin. Data jaetaan kahteen luokkaan, joita ovat kvalitatiivinen- ja kvantitatiivinen data. Kun puhutaan kvalitatiivisesta datasta, tarkoitetaan dataa, joka ei ole numeerista. Kvalitatiivisesta datasta puhuttaessa voidaan myös puhua laadullisesta datasta eli datasta, joka on arvioivaa sekä kuvaavaa.

Esimerkkinä kvalitatiivisesta datasta voidaan käyttää palautteenantoa. Palautteenannon ollessa suullista tai kirjallista, jossa ei ilmaannu numeroita, puhutaan kvalitatiivisesta datasta. Kvalitatiivisia esimerkkejä palautteenannosta:

- ”Järjestelmä toimi erittäin nopeasti”
- ”Ohjelmistosta puuttui viivakaaviot täysin”
- ”Ohjelmisto oli muuten hyvä, mutta värejä voitaisiin käyttää enemmän”

Kvalitatiivisen datan tunnistaminen on helppoa, sillä kyseessä on tieto, joka hyödyttää niin asiakasta kuin palvelun tuottajaa. Kvalitatiivinen analyysi onkin erittäin tärkeää juuri suunnitteluvaiheessa, kun halutaan tietää asiakkaiden tarpeista, haluista ja mielipiteistä enemmän. Laadullinen tieto koskee asiakkaan tunteita, käsityksiä ja haluja asioista, jotka asiakas kokee tarpeelliseksi. Nämä tiedot kerätään ylös ja niiden pohjalta tehdään myöhemmin päätöksiä ja ratkaisuja, unohtamatta kuitenkaan muuta kerättyä dataa kuten kvantitatiivista dataa. Kvalitatiivisen tutkimuksen kautta halutaan saada tietoa asiakkaasta, tuntemuksista, kokemuksista, motivaatiosta, yksittäisen henkilön omista mieltymyksistä ja haluista [13]. Kvalitatiivista dataa kerättäessä on suotavaa tietää asiakkaasta etukäteen joitakin asioita. Tämä helpottaa paitsi kysymysten laatimista haastateltaville, mutta auttaa myös ymmärtämään asiakasta ja hänen antamiaan vastauksia paremmin.

Kun on tiedossa, miten dataa lähdetään keräämään, täytyy tietää, kuinka kvalitatiivista dataa voidaan kerätä. Kvalitatiivista dataa voidaan kerätä vain hyvin rajatuin keinoin. Keruumenetelmät onkin rajattu usein näin:

- Henkilökohtainen haastattelu [14]
- Kohderyhmä haastattelu [14]
- Havainnointi ja tarkkailu [15]
- Tapaustutkimus (Case Study) [16]

Ennen haastattelun aloittamista haastateltavaa tulisi informoida kaikesta tarvittavasta. Tarvittava tieto sisältää asioita kuten missä tai miten haastattelu pidetään. On tärkeää myös ottaa huomioon eettiset periaatteet. Eritoten maininta siitä, kerätäänkö data anonyymisti ja luottamuksellisesti, on tärkeää [17]. On myös yleistä ja suotavaa mainita, mitä menetelmiä käytetään datan anonyymiyden säilyttämiseksi, jotta voidaan luoda luottamus haastateltavan kanssa. Tällä tavoin annetaan luotettava kuva itsestämme ja haastateltava tietää, kuinka hänestä kerättyä dataa käsitellään, säilytetään ja jalostetaan tulevaisuudessa. Ennen varsinaisen haastattelun aloittamista hyviin perustapoihin kuuluu kertoa, mihin haastattelu perustuu, mikäli tämä ei ole haastateltavalle selvinnyt etukäteen.[18]

Haastattelun sujuvuuden ja vastausten oikeellisuuden kannalta on otettava huomioon ympäristö, jossa haastattelu suoritetaan. Ympäristö ja sen luomat häiriötekijät vaikuttavat haastatteluun vääjäämättä, mutta riskiä voidaan kuitenkin pienentää. Riskiä voidaan pienentää hyvin yksinkertaisilla asioilla kuten varaamalla haastatteluun ajankohta, jolloin sekä tutkijalla että haastateltavalla on kunnolla aikaa paneutua itse haastatteluun. Häiriötekijöitä tulee välttää, mikäli se on mahdollista. Häiriötekijöihin lukeutuvat esimerkiksi erilaiset ylimääräiset äänet (puhelin, ihmiset), joten on suotavaa valita ympäristö, jossa on mahdollisimman vähän tilaa häiriötekijöille. Yleisesti ottaen haastattelut pyritään suorittamaan kerralla loppuun asti eli ylimääräisiä turhia keskeytyksiä pyritään välttämään, mikäli vain mahdollista. Valitettavasti haastattelut on usein sovittava haastateltavan aikataulujen mukaan. Tutkijan vastuulle jää häiriötekijöiden minimointi ja parhaaseen mahdolliseen tutkimusympäristöön mukautuminen sekä sopeutuminen kyseisessä tilanteessa.[19]

Kvalitatiivisen datan oikeellisuuden kannalta tutkijan tulee olla niin neutraali kuin mahdollista. Kysymykset tulee kysyä neutraaliin äänensävyyn, kehon liikkeiden tulee olla neutraaleja ja kaikki provosoivat teot ja eleet on jätettävä haastattelun ulkopuolelle. Tutkijan tulee osoittaa innostusta ja kiinnostusta aiheeseen ja haastateltavalta saatuihin vastauksiin [17]. Tämä saavutetaan helposti hymyilemällä ja nyökkäämällä eli olemalla läsnä haastateltavan kanssa mahdollisimman tavallisesti, mutta kuitenkin kiinnostavalla tavalla. Yksi tärkeimmistä asioista on välttää haastateltavan johdattamista, koska tällöin vastausten oikeellisuus saattaa vääristyä. Tutkijan ei tule esittää kysymystä muodossa “Oliko järjestelmä sinusta tarpeeksi nopea”, vaan vaihtoehtoisesti: “Mitä mieltä olit järjestelmän nopeudesta”, näin vältetään haastateltavan johdattamista ja varmistetaan vastausten oikeellisuus.[19, 18]

On hyvä tiedustella, mikäli haastateltavalla on tutkijalle kysymyksiä tai mietteitä haastatteluun liittyen. Haastattelun viimeisillä minuuteilla on vielä mahdollisuus saada haastateltavasta hieman lisää informaatiota irti vapaan keskustelun muodossa. Kvalitatiivisen datan keruumenetelmiä on muutamia, mutta kuitenkin rajallinen määrä. Henkilökohtainen haastattelu on vain yksi menetelmistä. Metodeja voidaan

kuitenkin soveltaa ja äsken mainitut asiat ovat asioita, jotka tulee ottaa huomioon lähes poikkeuksetta.[18]

2.4.1. Henkilökohtainen haastattelu

Henkilökohtainen haastattelu on yksi käytetyimmistä laadullisen datan keruumenetelmistä. Puhuttaessa henkilökohtaisesta haastattelusta tarkoitamme samalla tutkimushaastattelua. Tutkijana olemme halukkaita tutkimaan henkilöä ja henkilön omia mielipiteitä. Tutkimushaastattelutyypit jaetaan kolmeen eri tyyppiin:

- Jäsennelty
- Puoliksi-jäsennelty
- Järjestelemätön

Jäsennelty haastattelu on suullisesti hallittu kyselylomake. Tällöin haastateltavalta kysytään ennalta määrättyjä kysymyksiä luettelomaisesti. Yleistä tällöin on kysymysten kysyminen suoraan paperilta tai luettelosta, johon odotetaan suoraa ja tiivistettyä vastausta. Vastauksen pohjalta ei esitetä jatkokysymyksiä vaan kirjataan vastaus ylös ja jatketaan luettelomaisesti seuraavaan kysymykseen. Tämän kaltainen haastattelu on yleisesti helppo aikatauluttaa ja suhteellisen nopea suorittaa. Jäsennelty haastattelu on usein suhteellisen helppo toteuttaa olettaen että, tiedetään tarkkaan mitä halutaan kysyä. Kysymykset, siinä missä vastauksetkin, tulisivat olla ytimekkäitä, nopeita ja suoraviivaisia. Vastauksia ei ole tarkoitus jäädä pohtimaan yhdessä haastateltavan kanssa sen tarkemmin ja toisaalta syvällisempi tiedustelu jätetään pois kokonaan. Haluttaessa toteuttaa syvällisempää tiedustelua on parempi käyttää eri haastattelutyyppejä, kuten puoliksi-jäsenneltyä haastattelutyyppeä.[18]

Puoliksi-jäsennelty haastattelu, jota kutsutaan myös puolirakenteiseksi haastatteluksi, pitää sisällään pääkysymyksiä. Kysymysten tulisi olla rakenteeltaan samantapaisia kuin järjestetyssä haastattelussa, mutta haastateltavalle annetaan mahdollisuus syventyä aiheeseen yksityiskohtaisemmin tarpeen vaatiessa. Tämä tarkoittaa, että haastateltavalle annetaan tarkkojen kysymysten sijasta pikemminkin aihe, mistä puhutaan. Haastateltavalla ollessa enemmän asiantuntevuutta ja kokemusta juuri kysymykseen liittyen on mahdollista, että löydetään uutta tietoa, mitä ei ole pidetty oleellisena aikaisemmin, kun haastattelua on suunniteltu. Puoliksi-jäsennelty haastattelu onkin huomattavasti joustavampi kuin täysin jäsennelty haastattelu. On hyvin tavallista esittää jatkokysymyksiä haastateltavalle, mikäli vastaukset herättävät lisäkysymyksiä. Tämä vaatii kuitenkin haastateltavalta asiantuntevuutta sekä syvällisempää tietoa aiheeseen liittyen. Puoliksi-jäsennettyyn haastatteluun onkin suotavaa varata huomattavasti enemmän aikaa. Haastateltavalla ollessa enemmän asiantuntevuutta, kokemusta ja kiinnostusta kysymyksen aiheesta, on enemmän sääntö kuin poikkeus, että aiheesta poiketaan paikoittain jopa huomattavasti. Tämän seurauksena aikaa kuluu enemmän, mutta samalla tieto ja tietämys on huomattavasti syvällisempää.[18]

Järjestelemätön haastattelu on hyvä vaihtoehto silloin, kun tutkijalla ei ole tarkempaa tietoa siitä, mitä lähdetään kysymään. Tämän tapaisessa haastattelussa on yleistä, että ollaan vasta tutustumassa aiheeseen tarkemmin. Tutkijalla on yleisesti tällöin vain vähän tai lähes olemattomat ennakkotiedot aiheesta. Kyseisessä haastattelumuodossa on yleistä aloittaa kysymyksellä, joka on keskustelua avaava ja

todella laajaan aihealueeseen liittyvä. Kysymys voisi olla muotoa: “Kertoisitko hieman tietotekniikkataustastasi?”. Järjestelemätön haastattelu on juuri näiden avointen kysymysten takia erittäin aikaa vievä. Keskustelut voivat paikoittain poiketa aiheesta paljonkin, ja keskustelu voi edetä lähes mihin suuntaan tahansa. Haastattelutuokion pitäminen onkin kannattavaa silloin, kun tutkijalla ei ole aiheesta tietoa lähes lainkaan. Usein puoliksi-jäsennelty haastattelu on tehokkaampi, mutta joissain tilanteissa järjestelemätön haastattelu voi tuoda suurempaa tietämystä ja syvyyttä keskusteluun, kun haastateltavalle ei ole laadittu kunnolla kysymyksiä.[18]

2.4.2. Kohderyhmähaastattelu

Kohderyhmähaastattelussa on paikalla useita ihmisiä. Haastattelu suoritetaan ryhmässä eli kaikilta kohderyhmässä olevilta kysytään samaan aikaan kysymys. Aihe ja kysymykseen liittyvä keskustelu jatkaa tästä eteenpäin vapaamuotoisesti, mutta kuitenkin aiheeseen liittyen. Valitettavaa kohderyhmähaastattelussa on, että muiden vastaukset ohjaavat vääjäämättä keskustelua ja tämä vaikuttaa myös muiden ryhmässä olevien vastauksiin. Haastateltaessa ryhmää on siis tärkeää tiedostaa, että muiden vastaukset vaikuttavat myös muihin ryhmässä olevien vastauksiin. Tämän takia kohderyhmä haastattelussa haetaan usein vastausta hieman erilaisiin aiheisiin kuin henkilökohtaisessa haastattelussa.[18]

Haastattelun tavoitteena on usein saada kattavampi vastaus siitä, miksi jotakin tehdään juuri kyseisellä menetelmällä. Tällöin juuri hieman pohtiva ja usealta saatu vastaus on hyödyllinen, josta voidaan tehdä johtopäätöksiä myöhemmin. Mikäli tutkittavasta aiheesta ei ole tietoa, voidaan vaihtoehtoisesti suorittaa järjestämättömän haastattelun tilalla kohderyhmä haastattelu. Näin voidaan tutustua aiheeseen monen eri asiakkaan näkökulmasta samaan aikaan ja mahdollisesti saada suurempi kokonaiskuva tutkittavasta aiheesta. Kohderyhmähaastattelu on myös hyvä vaihtoehto eritoten silloin, kun halutaan tarkentaa tai laajentaa näkemystä tai vaihtoehtoisesti haastaa jokin muilla menetelmillä saatu informaatio.[18]

Tarkkaa määritelmää sille, kuinka suuri kohderyhmän tulisi olla, on vaikea määritellä. Kohderyhmät kuitenkin toimivat niin pieninä kuin suurina, kolmesta henkilöstä, jopa 14 henkilöön. Ryhmien koot tuovat omat haasteensa haastatteluun. Pienessä ryhmässä keskustelu voi jäädä hyvin vähäiseksi ja suppeaksi. Toisaalta suuremmassa ryhmässä keskustelu voi muuttua kaoottiseksi, sillä mielipiteitä ja henkilöitä on niin monta. Optimaalinen määrä henkilöitä kohderyhmässä on Stewartin mukaan kuuden ja kahdeksan henkilön välissä [20]. Samaa aikaan Stewart painottaa, että on parempi ottaa hieman liian monta henkilöä kohderyhmään kuin liian vähän, mahdollisten pois jäämisten takia [20]. Haastattelua valmistellessa Stewart antaa kaksi neuvoa. Kysymysten järjestyksen tulisi edetä tavanomaisista kysymyksistä tarkempiin ja yksityiskohtaisempiin kysymyksiin. Tärkeimmät kysymykset tulisi olla haastattelun loppupuolella, ja vähemmän tärkeät haastattelun alussa [20]. Valitettavasti näiden neuvojen noudattaminen juuri mainitulla tavalla on lähes mahdotonta, joten soveltaminen on erittäin suotavaa.

Tutkijan näkökulmasta haastattelu on yksinkertainen. Keskustelu aloitetaan kysymyksellä ja kysymyksestä aiheutuvaa keskustelua seurataan osallistumatta siihen. Tutkijan tulee pyrkiä pitämään keskustelu aiheessa, mutta ei kuitenkaan johdattaa keskustelua. Mahdollisimman laajan keskustelun aikaansaamiseksi, mikäli joku yksittäinen henkilö on liikaa äänessä, on tutkijan tehtävä pyrkiä antamaan muille ryhmässä enemmän valtaa omille mielipiteilleen. Näin keskustelu pysyy

monipuolisena ja antaa paremman kokonaiskuvan kysymykseen liittyvästä aiheesta. Hyvän haastattelun kannalta häiriötekijöiden minimointi on yhtä tärkeää kuin henkilökohtaisessa haastattelussa.[21]

2.4.3. Havainnointi ja tarkkailu

Kvalitatiivisen datan keräämiseen voidaan käyttää myös havainnointia ja tarkkailua. Tämän tyyppinen datankeruumenetelmä suoritetaan yleensä testihenkilöiden kannalta normaalissa ympäristössä. Voimme esimerkiksi tarkkailla, millä tavalla ihmiset käyttävät jotakin ohjelmaa. Testihenkilö voi olla joko tietoinen tai tietämätön siitä, että häntä tarkkaillaan, mutta on silti muistettava suorittaa tutkimus eettiset arvot huomioon ottaen. Tarkoituksena on tarkkailla henkilön tai usean henkilön käyttäytymistä päivittäisissä rutiineissa.[22]

Metodi on järjestelemätön, sillä tarkkailemme kuinka asiat toimivat ja edistyvät. Tutkimusmetodi on erittäin aikaa vievä ja sen seurauksena myös erittäin kallis toteuttaa. Tarkkailu- ja havainnointimenetelmällä on kuitenkin paikkansa kvalitatiivisen datan keruussa sen tiettyjen vahvuuksien takia. Tarkkailua suorittaessa, henkilöt toimivat normaalissa ja tutussa ympäristössä eli erittäin luonnollisesti. Tällöin voidaan havainnollistaa, kuinka henkilöt toimivat oikeasti juuri kyseisessä ympäristössä. Samaan aikaan voidaan luoda kokonaiskuva siitä, kuinka henkilöt suoriutuvat ja minkälaisia menetelmiä käyttäen. Havainnointi ja tarkkailu onkin erittäin hyödyllistä, kun halutaan joko haastaa tai tukea jotakin muuta väittämää, joka on saatu henkilökohtaista haastattelumetodia käyttäen. Mikäli asiat toimivat juuri kuten haastattelussa on väitetty, voidaan olla varmempia siitä, että asiat menevät juuri näin. Toisaalta, mikäli haastattelussa saatu tieto on ristiriidassa tarkkailussa havaitun tiedon kanssa, voidaan lähteä tutkimaan syitä ristiriidalle.[22]

Syitä ristiriidoille voi olla monia, mutta tarkkailussa ja havainnointimenetelmässä on omat haittapuolensa. Mikäli henkilöt, joita tarkkaillaan, ovat tietoisia tai huomaavat olevansa tarkkailun alla, heidän käyttöksensä voi muuttua. Tällöin henkilöt voivat mahdollisesti toimia eri tavalla, joka vääristää kerättävän tutkimusdatan luotettavuutta. Datan keruu tapahtuu usein tutkijoiden toimesta, joko kirjaamalla muistiinpanoja käsin, nauhoittamalla ääntä tai nauhoittamalla videomateriaalia. [22, 23]

Tutkijoilla voi jäädä asioita huomaamatta kirjaamisvaiheessa. Tällöin kaikkea dataa ei saada ylös ja jotain huomioonotettavaa voi jäädä pois tutkimusdatasta. Ääninauhoituksia ja videomateriaalia käytettäessä mahdollisuus tämänkaltaisen riskin toteutumiselle on pienempi, sillä nauhoitukset ja videomateriaali voidaan tutkia ja käydä läpi moneen kertaan. Ääninauhoituksissa on kuitenkin omat riskinsä sillä, ne voivat katketa kesken kaiken tai osa ääninauhoituksesta voi olla epämääräistä ja näin ollen sitä ei voida tulkita oikein. Videomateriaali voi myöskin kärsiä kuvauksen aikana samankaltaisista ongelmista.[22, 23]

2.4.4. Tapaustutkimus

Tapaustutkimuksen kannalta oleellista on avata keskeiset tiedot siitä, mihin osa-alueeseen tutkimus kohdistuu ja missä tutkimuksen todellinen painopiste sijaitsee. Ideana on, että kuka tahansa ulkopuolinen lukija tietää mihin tutkimuksella haetaan vastausta ja mihin lopputulokseen päästää tapauksen suhteen.[22, 24]

Tapaustutkimus sopii niin kvalitatiiviseen kuin kvantitatiivisen datan keruuseen. Menetelmä voi sisältää yhden tai useamman henkilön, mutta keskittyä myös tiettyyn ryhmään henkilöitä tai organisaatioon. Kyseistä menetelmää käytetään usein silloin, kun halutaan tutkia jotain tutkimusmielessä haastavaa tilannetta ja muuttujia on paljon. Tapaustutkimus loistaa tutkimusmenetelmänä, kun halutaan haastaa jokin vanha menettelytapa tai ajatusmalli.[24]

2.5. Kvantitatiivinen data ja analyysi

Puhuttaessa kvantitatiivisesta datasta tarkoitetaan datan arvoja, laskelmia tai numeroita. Kaikella kvantitatiivisella datalla on olemassa numeerisia arvoja, jotka ovat yksilöllisiä. Tietoja voidaan mitata ja käyttää hyväksi erinäköisissä laskelmissa ja tilastollisissa analyyseissä. Teetettyjen laskelmien ja analyysien pohjalta on mahdollista tehdä päätöksiä tosielämässä, kuten yritysmaailmassa. Kvantitatiivista dataa kuvataan myös määrällisenä datana. Datan tarkoitus on vastata kysymyksiin kuten: “Kuinka paljon?” ja “Kuinka monta?”.[25]

Tietojoukkojen kasvaessa parametrien avulla voidaan tehdä erilaisia analyysejä. Analyysejä hyödyntäen on mahdollista päätellä erilaisia laskennallisia asioita [25]. Esimerkiksi, kun vesimittarin kulutus ilmoitetaan numeerisena arvona ja tiedetään muita numeerisia arvoja, voidaan jalostaa tietoa kuvaamaan enemmän kuin kulutus tällä hetkellä. Tiedettäessä aika voidaan laskea muun muassa päiväkulutus, kuukausikulutus tai vuosikulutus. Hyödyntämällä kerättyjä tietoja voidaan tehdä arvioita seuraavalle vuodelle tai vaihtoehtoisesti päätellä veden hinta kuutiota kohden. Kuinka mittausparametrejä hyödynnetään ja liitetään toisiinsa, riippuu mihin kysymykseen halutaan vastaus. Kvantitatiivisen datan käsittelyssä on kolme vaihetta. Ensin tarvitaan mitattava aineisto, jota varten on suoritettava mittaukset. Toiseksi mittauksen perusteella tehdään analyysi kerätyistä mittauksista ja viimeiseksi muodostetaan johtopäätelmät. Kvantitatiivinen data on siitä helppo ja suosittu, sillä datan arvot ovat helposti hallittavissa erinäköisten matemaattisten menetelmien avulla.

2.5.1. Mittaus ja virheet

Mittauksia kerätessä käsin, koneellisesti tai jollakin muulla menetelmällä, keräämme muuttujia. Muuttujia voi olla lukemattomia erilaisia, mutta ne jakautuvat aina kahteen joukkoon, joita ovat itsenäinen muuttuja tai riippuvainen muuttuja. Riippuvainen muuttuja on aina riippuvainen itsenäisistä muuttujista [25]. Esimerkiksi vesihuollossa tutkittaessa painetta, lämpötila on itsenäinen muuttuja, jota kontrolloidaan. Paine on riippuvainen muuttuja, joka muuttuu lämpötilan muuttuessa ja on riippuvainen itsenäisestä muuttujasta.

Aina kun kerätään dataa eli esimerkiksi tässä tapauksessa mittauksia, on muistettava virheet. Virheiden aiheuttajia ovat ihmisen tekemä virhe, laitevirhe tai satunnainen virhe. Laitevirhe ja ihmisen tekemä virhe voidaan pyrkiä minimoimaan ja kontrolloimaan. Satunnaista virhettä kutsutaan virheeksi, jota ei voida ennakoida lainkaan eli ne ovat lähes aina väistämättömiä.

Ihmisen tekemä virhe tarkoittaa sitä, että jotain on tehty, mutta tämä ei ollut kyseisen henkilön tarkoitus. Tätä tapahtuu, kun suunnitelma on puutteellinen tai suoritus on riittämätön [26]. Esimerkiksi silloin, kun data on visualisoitu huonosti eikä vastaa haluttuun kysymykseen tarkoituksenmukaisesti, puhutaan ihmisen tekemästä

virheestä. Virheistä yksinkertaisin etenkin vesihuollon parissa on vedenkulutus. Vedenkulutus ilmoitetaan lähes poikkeuksetta vesihuollossa kuutioina (m^3). Mikäli käytetään pyöristyssääntöä eli 0.5 ja siitä ylöspäin pyöristetään ylöspäin ja muuten alaspäin, ei nähdä 490 litran ($0.49 \text{ m}^3 = 490 \text{ l}$) vedenkulutusta. Tarkasteltavan asian kannalta tämä voi olla erittäin merkittävää, silloin kun kyseessä on yksityishenkilö tai lähes merkityksetöntä, silloin kun kyseessä on käyttöpaikka.

Laitevirhe on yleinen ja tunnistetaan vesihuollon ja laitevalmistajien toimesta. Sama laite voi antaa eri mittauksia. Aiheuttajia voi olla esimerkiksi laitteen ikä, kalibrointi tai mittauspaikka. Laitevalmistajat antavat tuotteen mukana pitkän listan mahdollisista virheistä, jotka voivat toteutua laitteessa. Lisätietoa haluttaessa voidaan ottaa valmistajaan yhteyttä ja kysyä lisätietoja [27]. Data-analyysin kannalta onkin tärkeää tiedostaa kaikki mahdolliset virheet ja ottaa ne huomioon analyysiä tehdessä välittämättä siitä, tekeekö analyysin tietokone vai henkilö.

2.6. Visualisoituminen käsitteenä

Visualisointi on tärkeää ja sille on tarvetta niin nyt kuin tulevaisuudessakin. Visualisointi voidaan toteuttaa erilaisia työkaluja käyttäen, mutta toteutuksessa on kiinnitettävä huomioita erilaisiin seikkoihin. Visualisoinnin avulla pystytään parhaimmillaan jalostamaan data informaatioksi, tiedoksi ja tietämykseksi.[28]

Datan visualisoinnin hyöty kantaa juurensa ihmisen erittäin tehokkaasta havainnointikyvystä, eritoten silloin, kun puhutaan näöstä. Näön avulla ihminen kykenee hahmottamaan ilmiöitä sekä säännöllisyyttä visualisoidusta datasta, joka muissa havainnointimenetelmissä saattaisi jäädä huomaamatta.[29]

Olemme aikakaudella, jossa dataa kerätään koko ajan enemmän ja enemmän. Etäluettaviin mittareiden määrä kasvaa ja reaaliaikaisen datan kerääminen tuottaa valtavia määriä dataa. Vuonna 2015 IBM julkaisi raportin, jonka mukaan 90% kaikesta datasta on kasaantunut pelkästään kahden edellisen vuoden aikana (2013-2015). On tärkeää liiketoiminnan kannalta, että kerätyn datan avulla voidaan johtaa ja tehdä päätöksiä entistä nopeammin ja tarkemmin. Kerätyn datan määrä on kasvamaan päin myös vesihuollon puolella. Tämä tarkoittaa väistämättä myös sitä, että datan tulkitseminen on hankalampaa ja lopulta, datamäärien kasvaessa, lähes mahdotonta tehdä manuaalisesti. Datan visualisointi on osoitettu tehokkaaksi menetelmäksi näyttää kriittisimmät ja tärkeimmät asiat suuresta määrästä dataa. Tutkijat ovat myös havainneet, että värilliset visuaaliset esitykset lisäävät halukkuutta lukea jopa 80% [30]. Artikkelissa samat tutkijat mainitsivat myös, että visuaalisesta esityksestä voidaan muodostaa käsitys jo kymmenesosa sekunnissa.[30]

Massadatan muuttaminen informatiiviseksi toteutetaan käyttämällä tietokoneita ja ohjelmien tuomaa analytiikkaa. Visuaalisesta esityksestä voidaan tehdä analyysijä nopeammin ja poimia halutut ydinasiat datasta.[31] Dataa ei kannata kuitenkaan visualisoida, ellei se ole järkevää. Muistisääntönä datan visualisoinnille kannattaa käyttää lausetta: "Mikäli asia voidaan kuvata selvästi sanojen avulla, visualisointi ei todennäköisesti ole tarpeellista." Perusajatus visualisoinnissa ja sen tuomalla tehokkuudella on tuottaa ymmärrettävää tietoa katsojalle. Esityksestä lukijan tulee pystyä jalostamaan vastauksia tai muita hyötyjä itselleen tai toiselle riippuen aiheesta.[32]

Datan visualisoinnissa on onnistuttu erityisen hyvin silloin, kun lukija ymmärtää kuvan tai diagrammin, mitä se esittää ja pystyy tuottamaan sen avulla vastauksia taikka tekemään päätelmiä. Mikäli visualisointi on sekavaa, johtuu se usein väärin valituista

visualisointimenetelmistä ja niiden toteutuksesta. Kaikki edellä mainitut kriteerit on otettava huomioon, kun suunnitellaan näytettäviä taulukoita tai diagrammeja.[32]

2.7. Visuaalinen tiedustelu

Käsitys visuaalisten asioiden ymmärtämisestä ja tarkasta näkemyksestä on aivan toisenlainen kuin kuvittelemme. Perusajatus, että ihminen ymmärtää näkemänsä ensisilmäyksellä, ei pidä paikkansa. Hereillä ollessa ihmisen ympärillä tapahtuu lukemattomia asioita, ja todellisuudessa emme kiinnitä huomioita kuin murto-osaan kaikesta tapahtuvasta. Kriittinen informaatio, johon kuitenkin kiinnitämme huomioita, riittää kantamaan ihmistä eteenpäin ja suorittamaan päivän tehtävät, ainakin jossakin määrin.[33]

Esimerkiksi päivän aikana kävellessä kadulla ihminen näkee lukemattomia kasvoja. Näistä kaikista kasvoista ihminen pystyy muistamaan vain muutaman, ja nämä kasvot ovat sellaisia, joihin kiinnitämme erityisen paljon huomiota. Eräässä tieteellisessä tutkimuksessa, Daniel Simons ja Daniel Levin suorittivat kokeellisen tutkimuksen, jossa tutkittiin mitä ihminen huomaa [33]. Tutkimuksessa tutkimusryhmän henkilö kysyi neuvoa henkilöiltä (apuna käytettiin karttaa, jotta henkilö kiinnittää karttaan huomionsa hetkeksi). Kesken kaiken tutkimusryhmän henkilö vaihdettiin toiseen niin, että karttaa tutkiskeleva henkilö ei huomannut. Tutkimusryhmän henkilöllä saattoi olla eri vaatetus, eri hiusten väri tai eri sukupuoli kuin aikaisemmin. Henkilöt, joilta neuvoja reittiohjeisiin oli kysytty, eivät huomanneet muutoksia. Kun tutkittiin hiusten, vaatteiden ja henkilön vaihdosta, yli 50% ei huomannut muutosta lainkaan. Tämä kuvaa jo hyvin kuinka näön avulla tehty havainnollistaminen tapahtuu vain, kun itse kiinnitämme siihen huomiota. Tapaus, jossa testiryhmän henkilön sukupuoli vaihtui kesken kaiken, ei aiheuttanut joissakin tapauksissa mitään epäilyä tai hämmennystä henkilöissä, mutta tämä oli kuitenkin harvinaisempaa. Selitys tälle löytyy ihmisestä ja sen aivokapasiteetin maksimaalisesta hyödyntämisestä. Kyseisessä tilanteessa aivot ovat käsitelleet kaiken näön avulla havaitut asiat, mutta aivot eivät ole tulkinneet niitä relevanteiksi tai tärkeiksi. Ihmisen vaihtuminen ei ole relevantti tai tärkeä asia, kun keskitytään kartan tutkimiseen ja parhaiden neuvojen antamiseen. Tällä tavoin aivot optimoivat energian käyttämistä ja minimoivat energian kulutuksen.[33]

Tiedettäessä, kuinka ihmisen aivot käyttäytyvät erilaisissa tilanteissa, voidaan hyödyntää sen tuomaa tietoa myös graafisessa suunnittelussa. Kun käytetään hyväksi aivoissa tapahtuvaa prosessia, voidaan kuvien kautta nopeuttaa ja helpottaa oikean asian löytämistä. Silmä on työkalu, jolla ihminen pystyy näkemään asian ja aivot tekevät työn havainnon muuttamisesta informaatioksi. Suuresta massasta visuaalisia havaintoja aivot keskittyvät ensin siihen mikä on relevanttia. Tapauksessa on oleellista tietää, mitä aivot tiedustelevat ja mitä ne etsivät. Erityisesti tottumukset ja tutuksi tulleet visuaaliset menetelmät ohjaavat aivojamme etsimään juuri tiettyjä asioita. Erityisesti kohderyhmän tunnistaminen ja heidän käyttämät menetelmät ovat tärkeitä suunniteltaessa visuaalisia kaavioita.[33]

2.8. Selkeä visualisointi

Visualisoinnissa on otettava itse visuaalisen asian lisäksi huomioon kohderyhmä, jolle visualisoitu tuote on suunnitteilla. On pystyttävä vastaamaan kysymyksiin: kuka on käyttäjä, mihin tarkoitukseen visuaalista esitystä käytetään ja mikä on käyttäjän

asiantuntemus aiheeseen liittyen. Suunnittelussa sekä toteutuksessa on oleellista tiedostaa, että eri käyttäjäryhmillä on erilaiset tarpeet ja näihin tarpeisiin on pystyttävä vastaamaan. Johtajille tarvitaan erilaiset näkymät, kuin projektipäällikölle. Hierarkiassa alemmalla tasolla olevan työntekijän kannalta projektipäällikölle tarjottu visuaalisesti muotoiltu näkymä on usein turha. Näkymä tarjoaa oleellista tietoa juuri projektipäällikölle, kun sen pitäisi tarjota sitä alemman tason työntekijälle. Selkeän visualisoinnin kannalta tärkeintä on muistaa vastata kysymykseen: “Mihin kysymykseen halutaan löytää vastaus datan avulla “[30]

Tutkittaessa kohderyhmää ja kohderyhmän vaatimuksia selkeään visualisointiin, täytyy ottaa monia asioita huomioon. Huomioon on otettava muun muassa ikä, kulttuuri ja värit, joita visualisoinnissa käytetään. Vanhemmille ikäryhmille suurempi fonttikoko on suotavaa, kun nuoremmille sillä ei välttämättä ole merkitystä. Vanhempi ikäpolvi ei yleisesti ottaen ole samalla tasolla nuorempien kanssa, kun puhutaan asiantuntevuudesta digitaalisten palveluiden parissa, kuten niiden toiminnallisuudessa. Värien käytössä on tärkeää muistaa käyttää värejä, jotka ovat selvästi erotettavissa toisistaan. Mikäli värit eivät ole helposti havaittavissa, värisokeat käyttäjät kärsivät, sillä he eivät välttämättä erota haluttuja asioita esityksestä.[34]

Kulttuuri tuo omat haasteet joita ovat muun muassa lukusuunta ja palvelussa käytettävät kielet. Lukusuunta ja siihen liittyvät ongelmat ovat ajankohtaisia, kun palvelua tarjotaan ympäri maailmaa. Kieli ja kielivalinnat ovat tärkeä osa palvelua. On tutkittava käyttäjäryhmiä ja heidän käyttämiään kieliä eri palveluissa. Kulttuuriset erot on otettava huomioon kokonaisuudessaan suunnittelussa ja tuotteistamisessa. Teknologian kehitys tuo myös muita haasteita, kuten palvelun toiminnallisuus erilaisilla laitteilla tabletilla, mobiililaitteilla sekä tietokoneella. Selkeän visualisoinnin kannalta kaikki edellä mainitut asiat on syytä ottaa huomioon.[34]

Datan visualisointi tuo vielä omat haasteensa, mutta ottamalla huomioon ainakin viisi ominaisuutta, saadaan visualisointi hyvälle tasolle. Nämä ominaisuudet ovat ymmärrettävyys, selkeys, luotettavuus, tuttavallisuus ja miellyttävyys. Ymmärrettävyydellä tarkoitetaan sitä, että visuaalisessa esityksessä ei tule olla mitään ylimääräistä, vaan vain ja ainoastaan tarvittava informaatio. Selkeys liittyy paljolti ymmärrettävyyteen, mikä tarkoittaa, että on tärkeää panostaa informaation esitykseen niin, että asiakas ymmärtää sen nopeasti. Luotettavuus saavutetaan silloin, kun käyttäjälle näytettävä informaatio on selkeää, ja käyttäjä voi luottaa näytettävän informaation oikeellisuuteen. Tuttavallisuudella tarkoitetaan esittämistapoja, jotka ovat ennestään tuttuja käyttäjäryhmälle. Tutustuminen käyttäjäryhmään ja heidän käyttämiinsä ja totuttuihin visuaalisiin tapoihin näyttää informaatiota on suotavaa. Tällä tavoin voidaan välttää turhat sekaannukset. Viimeisenä mielekkyys tarkoittaa, että tarjotun visuaalisen esityksen tulee olla mielekäs tulkita, lukea sekä analysoida. Kaikki mainitut ominaisuudet huomioon ottamalla saadaan toteutettua hyvä visuaalinen esitys asiakkaalle.[34]

2.9. Visualisoinnin tavoite

On lähes mahdotonta ymmärtää massadataa, mikäli sitä ei ole “pakattu” muotoon, josta ihminen pystyy sitä tulkitsemaan kohtuullisessa ajassa. Visuaalinen ilmaiseminen on tällöin erittäin käytetty tapa, jolla voidaan näyttää ja korostaa tiettyjä asioita massadatasta. Visualisoinnissa on otettava huomioon käyttäjäryhmät, jotka

tulkitsevat informaatiota, joka on näkyvissä. Informaation tulee olla ymmärrettävissä juuri käyttäjäryhmälle – ei sen tekijälle tai muulle taholle.

Ennen työkalujen valitsemista on tutkittava mitä käyttäjäryhmä haluaa nähdä. Visuaalisen esityksen tulisi olla sellainen, joka kuvaa käyttäjälle juuri halutun asian eikä mitään ylimääräistä. On tapauksia, jolloin halutaan itse tehdä päätelmiä visuaalisesta esityksestä ja toisaalta välillä halutaan tarkempaa tietoa kysytyyn asiaan. Tämän tunnustamisessa auttaa käyttäjäryhmä tunteminen, mikä tarkoittaa sitä, että tiedetään tarkasti, halutaanko vain suora vastaus vaiko hieman enemmän, jolloin tehdään saadusta informaatiosta lisäanalyysi itse. Informaation helpompaa tulkittamista auttaa työkalut, kuten datan järjestäminen aakkosjärjestykseen tai vaihtoehtoisesti aikajärjestykseen. Visualisoinnin päätavoitteena on antaa käyttäjälle nopeasti tulkittava ja ymmärrettävä esitys, joka tuottaa käyttäjän haluaman vastauksen.[34]

2.10. Työkalut visualisointiin

Datan visualisointiin käytetään työkaluja, jotka helpottavat kuvaamaan tiettyjä asioita nopeasti sekä tulkitsemaan niitä helposti. Koponen & Hildén esittävät kirjassaan Data visualization handbook kaaviota, joiden käyttö on yleistä ja samaan aikaan nopeasti ja helposti ymmärrettävissä, useille eri käyttäjäryhmille [32]. Kaavioita ovat muun muassa pylväskaavio ja sen variaatiot, viivakaavio, piirakka- ja donitsikaavio sekä tutkakaavio ja datakartat. Tämän työn ohessa kaavioista käytetään myös nimitystä diagrammi, joka tarkoittaa samaa asiaa.[32]

Kaaviota käytettäessä on tiedostettava, mitä kaaviota minkäkin asian esittämiseen tulisi käyttää. Eri asioiden selkeään näyttämiseen osa kaaviosta soveltuu paremmin kuin toiset – toisaalta saman asian esittämiseen voidaan käyttää useampaa eri kaaviotyyppiä. Muutamia yleisesti tunnettuja ja käytettyjä kaaviotyyppisiä, kun halutaan kuvata:

- Muutosta ajan suhteen
 - Pylväskaavio
 - Viivakaavio
 - Bullet-kaavio
- Vertailu
 - Pylväskaavio
 - Waterfall eli vesiputouskaavio
- Jakaantuminen
 - Tutkakaavio
 - Box-Plot eli laatikkokaavio
- Kokonaisuus ja osakokonaisuus
 - Pylväskaavio
 - Piirakkakaavio
 - Pinottu pylväskaavio
- Sijoitus
 - Pylväskaavio

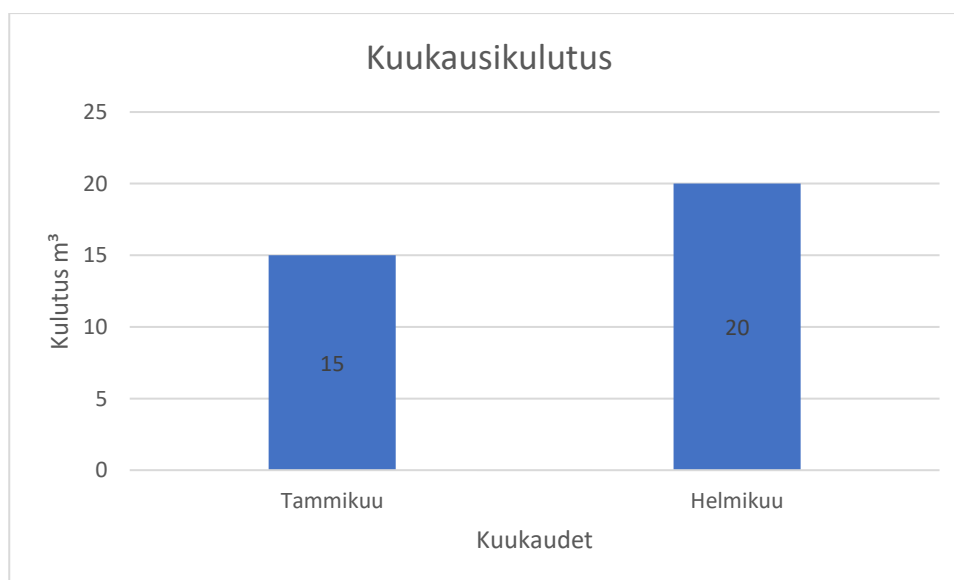
Kuten huomataan, samat kaaviot toistuvat useassa eri kategoriassa. Tämä johtuu pääosin siitä, että kyseisiä kaavioita käytetään yleisesti niin tieteessä kuin yritysmaailmassa. Mainitut kaaviot on todettu myös helposti ja nopeasti

ymmärrettäviksi. Vähemmän käytettyjä kaavioita on Bullet-kaavio, vesiputouskaavio ja laatikkokaavio. Kun vähemmän käytettyjä kaavioita käytetään, käyttäjäryhmän on oltava tietoinen, kuinka kaavioita tulkitaan. Vähemmän käytettyjen kaavioiden tulkinnan opettamiseen tarvitaan aikaa ja resursseja, niin palvelun tuottajalta kuin kuluttajan puolelta. Väärin ymmärretty kaavio voi aiheuttaa kohtalokasta vahinkoa ja sen takia yleisesti käytettyjen kaavioiden käyttäminen on turvallisempaa.[30]

2.10.1. Pylväskaavio

Pylväskaaviossa on tiedettävä, mitä arvoja sijoitetaan x-akselille ja mitä y-akselille. Mistä arvosta aloitetaan x- sekä y-akseli ja miksi. Pylväskaaviota käytettäessä on hyvä muistaa, että se voidaan esittää sekä pystysuorassa että vaakasuorassa. Kuten huomataan kuvasta 2 ja 3, pystysuorassa pylväsdigrammissa numeeriset arvot on yleisesti sijoitettu y-akselilla. Numeeristen arvojen ollessa y-akselilla, x-akselille sijoitetaan tarkasteltava asia. Ylimääräiset asiat, kuten värien merkitys kerrotaan usein erikseen, jotta katsojalle ei jää epäselvyyttä mitä eri värit kuvaavat. Kun pylväskaavio kaadetaan vaakatasoon, akseleiden suuret vaihtavat paikkaa lähes poikkeuksetta. Tällöin y-akselilla on tarkasteltavat asiat ja x-akselilla numeeriset arvot.[32]

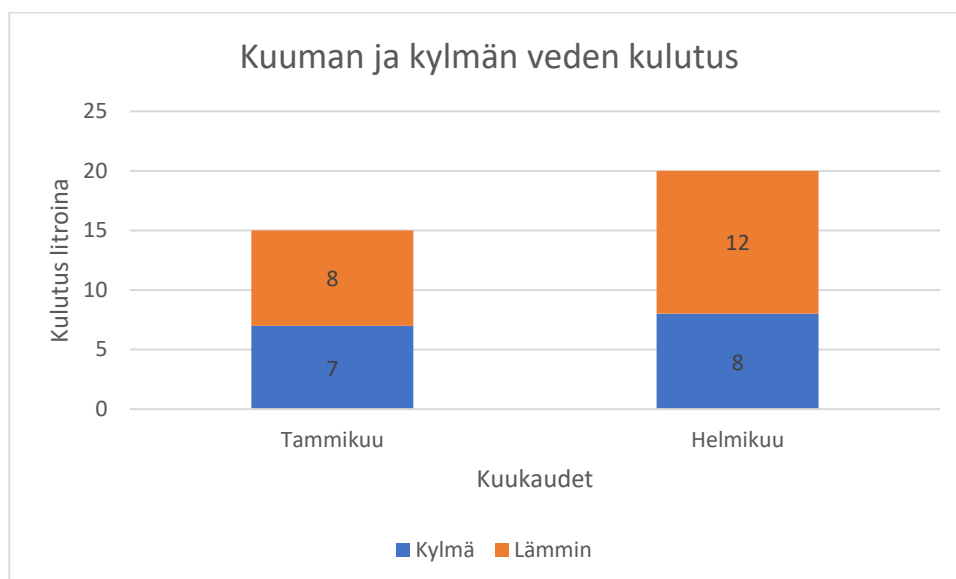
Pylväs- tai toiselta nimeltä palkkidiagrammit(vaakatasossa) on erittäin käytetty visualisoinnin menetelmä. Pylväskaaviosta on myös useita variaatioita kuten pinottu pylväskaavio. Pinotussa pylväsdigrammissa on hyödyllistä käyttää värejä, jotka näkyvät myös selvästi kuvassa 3. Kun puhutaan pituudesta, liitetään se yleensä laskuun tai nousuun. Tätä käytetään hyväksi usein muun muassa pylväsdigrammeissa. Esimerkiksi, vedenkulutuksen ollessa suurempaa kuin edellisenä kuukautena, palkit ovat eripituisia. Kulutuksen ollessa suurempaa palkki on pitempi ja toisaalta lyhyempi, kun kulutus on pienempää. Kyseinen ero voidaan nähdä hyvin kuvasta 3, jossa palkki 1 kuvaa pienempää kulutusta kuin palkki 2.[32]



Kuva 2. Kulutus kahtena eri kuukautena.

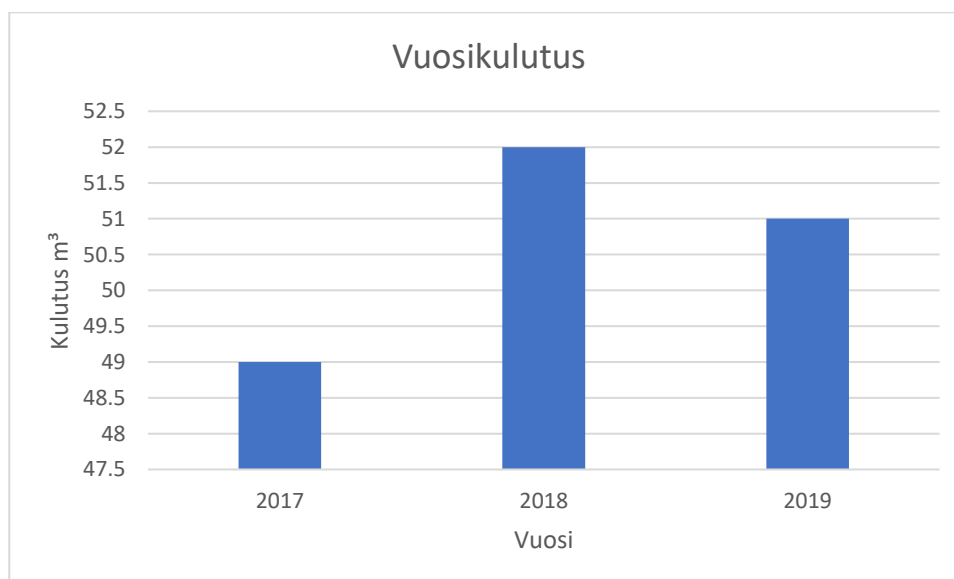
Kategorioita voidaan erotella pituuden lisäksi myös väreillä. Samassa diagrammissa tai jopa samassa parrussa voidaan näyttää kahden eri kategorian arvot

samassa kaaviossa. Kuvassa 3 kokonaiskulutus on palkissa olevien numeroiden summa. Tässä tapauksessa tammikuussa, lämpimän osuus on kahdeksan, kylmän seitsemän ja summa 15.[32]



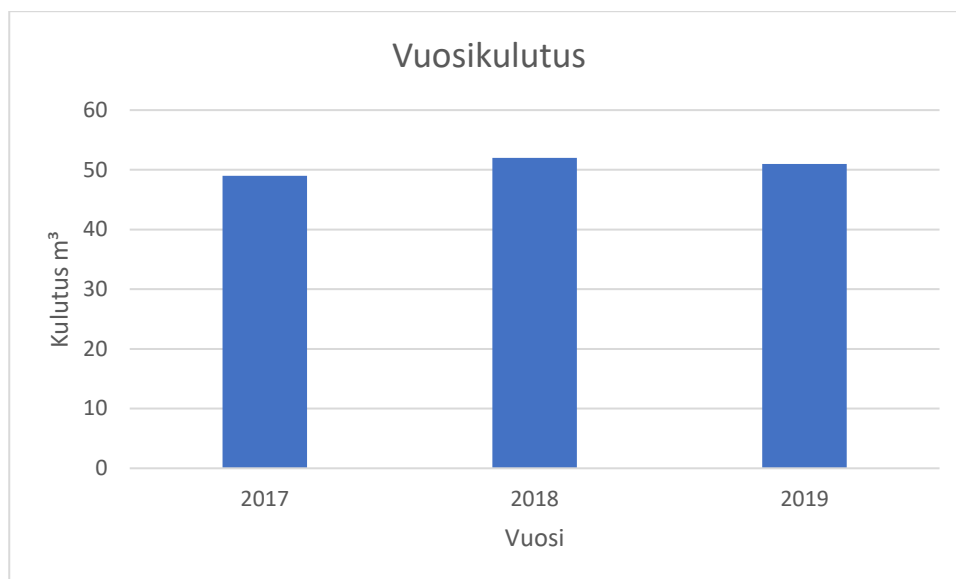
Kuva 3. Vedenkulutuksen jakauma.

Vaikka pylväsdiagrammit ovat yleinen tapa esittää suurempaa määrää dataa kuvamuodossa, siinäkin voidaan tehdä suuria virheitä. Hyöty pylväsdiagrammista häviää, mikäli tarkasteltavaa asiaa ei pystytä tulkitsemaan oikein. Esimerkiksi, kun tarkastellaan vedenkulutusta, halutaan nähdä suuret poikkeamat normaalissa kulutuksessa. Tämän avulla voidaan havaita mahdollisia vuotoja tai muita ongelmia. Kun datan visualisointi eli kuvaus on esitetty kuin kuvassa 4, voi tapahtua virheitä (y-akseli kuvaa vedenkulutusta kuutioina, x-akseli lukemia vuotuisia lukemia). Nopeasti tulkittaessa näyttää, että vedenkulutuksessa olisi suuria eroja (y-akseli puolen yksikön välein) ja kohteessa olisi mahdollisesti vuoto tai muita ongelmia.



Kuva 4. Y-akselin aloitus 47,5.

Toisaalta kuva 5, jossa y-akselin muuttuja on sama, mutta tarkkuus eri, (kymmenen yksikön välein) kuvaa samaa tilannetta huomattavasti paremmin. Heitot kulutuksessa näyttävät olevan lähes olemattomia ja nopean tulkinnan jälkeen voidaan päätellä kaiken toimivan normaalisti.

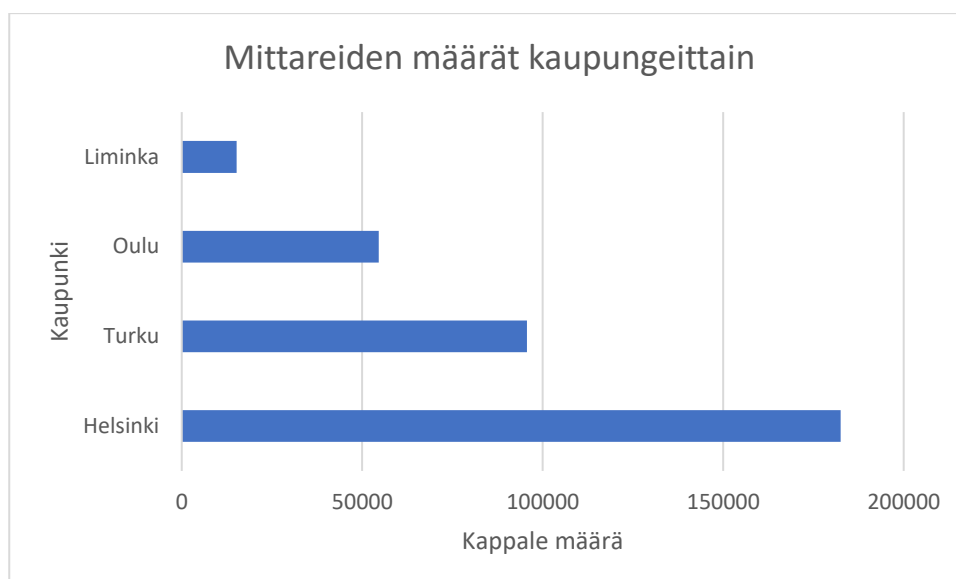


Kuva 5. Y-akselin aloitus nolasta.

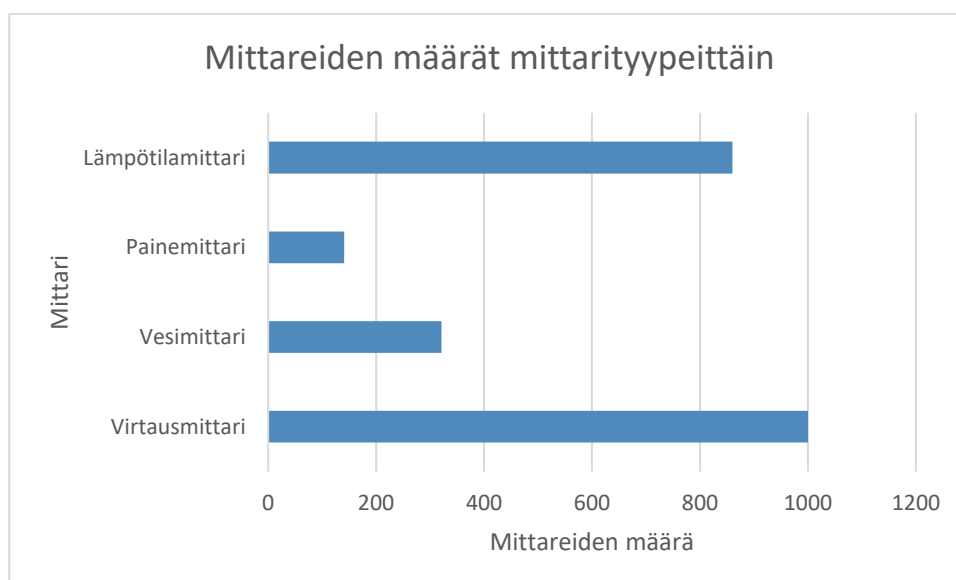
Vesihuollossa jopa muutamien kymmenien kuutioiden heitto on normaalia riippuen mitattavasta kohteesta. On tärkeää, että esitettävät suureet ovat oikeassa mittasuhteessa ja oikein määritelty mitattavaa ja tiedusteltavaa asiaa kohden [2]. Näin saadaan optimaalinen ja maksimaalinen hyöty irti visualisoinnista ja mahdolliset analyysit voidaan tehdä oikein ja luotettavasti. Ongelmia juuri kulutuksen seurannassa on havaittu aikaisemmin suunniteltaessa Vesitiedon palvelua Vesitiliä. Vesitilissä näytetään vedenkulutusta asiakkaille niin litroina kuin kuutiona. Ongelmatilanteet syntyivät, kun asiakkaalle näytettiin pylväskaaviossa kulutus litroina. Mikäli vuotuinen vedenkulutus vaihtelee esimerkiksi 20 kuutiota, litroina ero tarkoittaa 20000 litraa.[2]

Tarkasteltaessa yhä suurempaa kokonaisuutta, kuten mihin vuodenaikaan tai mihin kellon aikaan kulutus on suurinta yleisellä tasolla, tarvitaan uusia menetelmiä visualisointiin. Richard Saul Wurman on esitellyt vuonna 1989 viisi tapaa järjestää informaatiota [35]. Nämä viisi tapaa ovat kategoria, aika, sijainti, hierarkia ja aakkosjärjestys, joista käytetään lyhennettä LATCH (Location, Alphabet, Time, Category, Hierarchy).[32]

Vesimittareiden tai yleisesti mittareiden määrää tarkastellessa alueittain on hyödyllistä järjestää data sijainnin mukaan. Kuvasta 6 voidaan nähdä missä kaupungeissa mittareita on paljon ja missä taas vähän (mittareiden lukumäärä on mielikuvituksellinen) Tuotetun informaation avulla voidaan tehdä päätelmiä kuten tulisiko mittari kantaa mahdollisesti lisätä, jotta pysytään kilpailevassa asemassa.[32]



Kuva 6. Mittareiden määrä kaupungeissa (määrät ovat mielikuvituksellisia eivätkä todellisia!).



Kuva 7. Mittarit aakkosjärjestyksessä (määrät ovat mielikuvituksellisia eivätkä todellisia!).

Aakkos-järjestelmän käyttäminen kuten kuvassa 7 on hyödyllistä, kun halutaan tarkastella mittareita järjestelmällisesti tai alueellisesti. Myös muut järjestysmenetelmät kuten hierarkia ja kategoria ovat käteviä, mutta käytetyin on aikajärjestys. Aikajärjestystä käytetään paljon, kun muutetaan dataa informatiiviseksi. Mikäli aikaleima (Timestamp) löytyy useammasta eri asiasta, kuten mittarin valmistuksesta, mittarin asennuksesta ja lukeman lähetyksestä, voidaan kaikkia näitä kategorioita tarkastella erikseen ja järjestää aikajärjestykseen kategorian mukaan. Mittarin valmistuksen aikaleiman tarkasteleminen voi olla hyödyllistä, kun mittarierässä on esimerkiksi valmistusvirheitä. Näin löydetään mahdolliset virheelliset mittarit nopeasti ja voidaan tehdä toimenpiteitä. Mittarin asennus kertoo vesihuoltolaitokselle, milloin mittari on asennettu uudelle käyttöpaikalle ja milloin voidaan odottaa mittarin toimivan. Toisaalta tieto kertoo, milloin mittarin omistajaa

voidaan alkaa laskuttamaan. Lukeman lähetyksestä voidaan nähdä vedenkulutus ja mahdollisesti muuta dataa ja tarkastella mittarin toimintoja ja datan oikeellisuutta. Pelkästään datan järjestely analysointina tekee tiedosta informatiivista ja luettavaa – joskin myös muita menetelmiä on hyödyllistä käyttää.[2, 32]

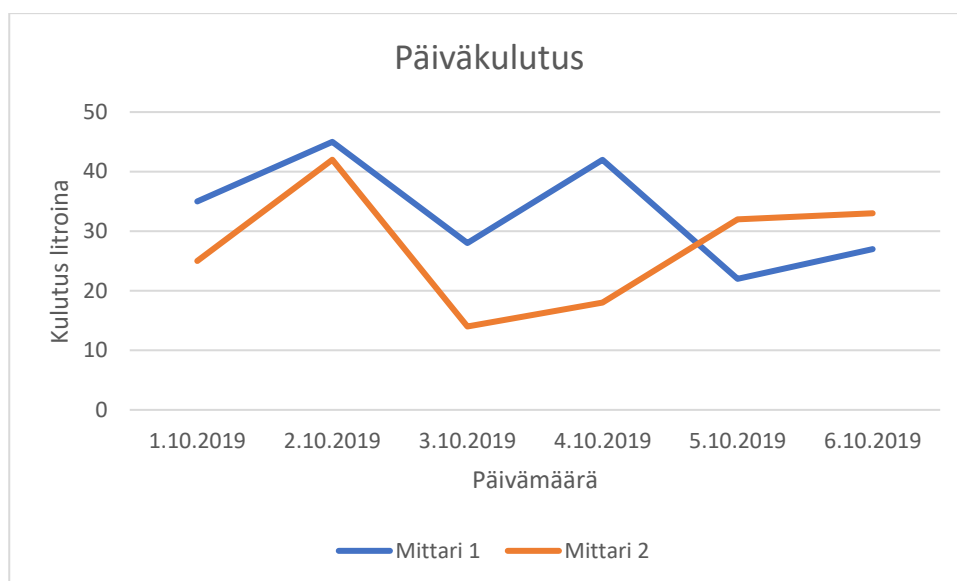
2.10.2. Viivakaavio

Ajan muutosta kuvattaessa käytetään usein viivakaaviota. Viivakaaviosta erityisen hyvän tekee sen helppolukuisuus ja muutosten näkeminen. Viivakaaviossa x-akseli kuvaa aikaa ja y-akseli keskittyy määrään. Kuvasta 8 voidaan nähdä, kuinka selvästi muutokset ovat nähtävissä viivakaaviossa.



Kuva 8. Viivakaavio, määrän muuttuminen ajan funktiona (määrät ovat mielikuvituksellisia eivätkä todellisia!).

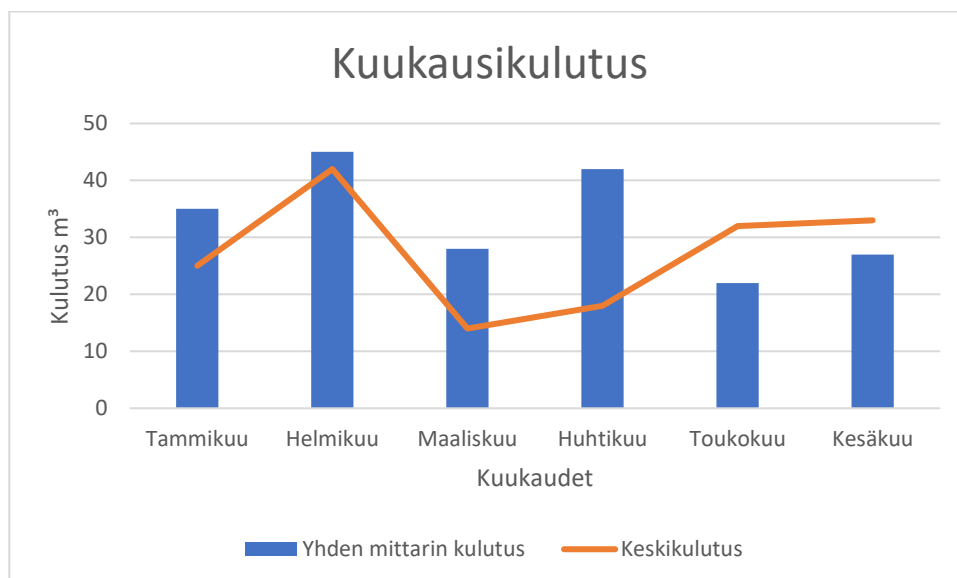
Viivakaavio käsittelee puuttuvia arvoja hyvin, sillä se vetää viivan pisteestä seuraavaan pisteeseen välittämättä puuttuvista pisteistä. Näin ollen, mikäli jokin arvo puuttuu datasta, kaavio ei välttämättä kuvaa käyrää täysin oikein, mutta trendi on kuitenkin suuntaa antava. Viivakaavio on sitä tarkempi mitä useampaa datapistettä ollaan kuvaamassa. Viivakaaviossa voidaan kuvata myös useampaa ajan funktiota samaan aikaan. Tällöin voidaan tehdä arviointia kahden funktion välillä, kuten tarkastellessa vedenkulutusta kahden mittarin välillä (katso kuva 9).[32]



Kuva 9. Viivakaavio, kahden mittarin lukeman vaihtelu ajan funktiona (määrät ovat mielikuvituksellisia eivätkä todellisia!).

2.10.3. Pylväs- Viivakaavio

Yhdistettäessä pylväs- ja viivakaavio voidaan esittää kaksi mittausta samassa visuaalisessa esityksessä. Tällöin ensimmäinen mittaus on esitetty esimerkiksi pylväskaaviomuodossa ja toinen mittaus piirretty pylväiden päälle viivakaaviona (kuva 10).



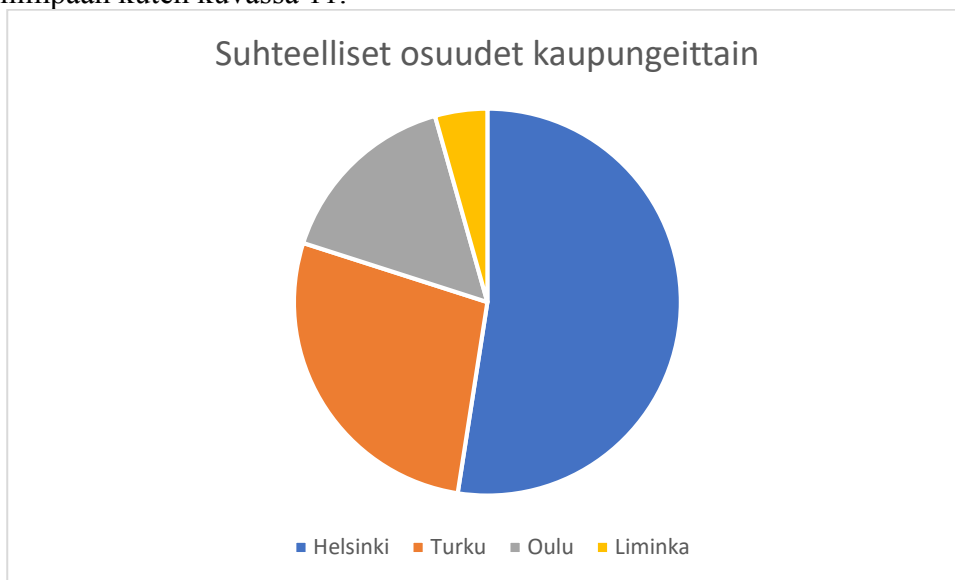
Kuva 10. Pylväs- viivakaavio (määrät ovat mielikuvituksellisia eivätkä todellisia!).

Kuvassa 10 pylväskaavio kuvaa mielikuvituksellista kulutusta yhdelle mittarille vuonna 2019 kuuden kuukauden ajan. Kuvioon on myös lisätty keskikulutus (kaikkien

mittareiden kulutuksen summa jaettuna mittareiden määrällä) viivakaaviota käyttäen näyttämään toinen mittausta. Kuvasta 10 on helppo ja yksinkertainen katsoa onko kulutus tiettyinä kuukausina suurempaa vaiko pienempää kesikukulutukseen verrattuna.[30]

2.10.4. Piirakkakaavio

Piirakkakaaviosta käytetään usein myös nimitystä ympyrädiagrammi. Ympyrädiagrammissa jaetaan tietyt osuudet omiin sektoreihin. Nämä osuudet ovat suhteellisia ja niitä kuvataan usein prosentteina. Piirakkakaavion avulla pystytään vertailemaan osuuksia tietyistä kokonaisuudesta. Ympyrädiagrammia käytetään yrityksissä, sillä se on helppo hahmottaa. Tieteellisissä julkaisuissa ympyrädiagrammia ei usein käytetä, sillä se kuvaa erittäin huonosti eroja osuuksien välillä. Tyypillistä ympyrädiagrammeille on järjestää osuudet suurimmasta pienimpään kuten kuvassa 11.



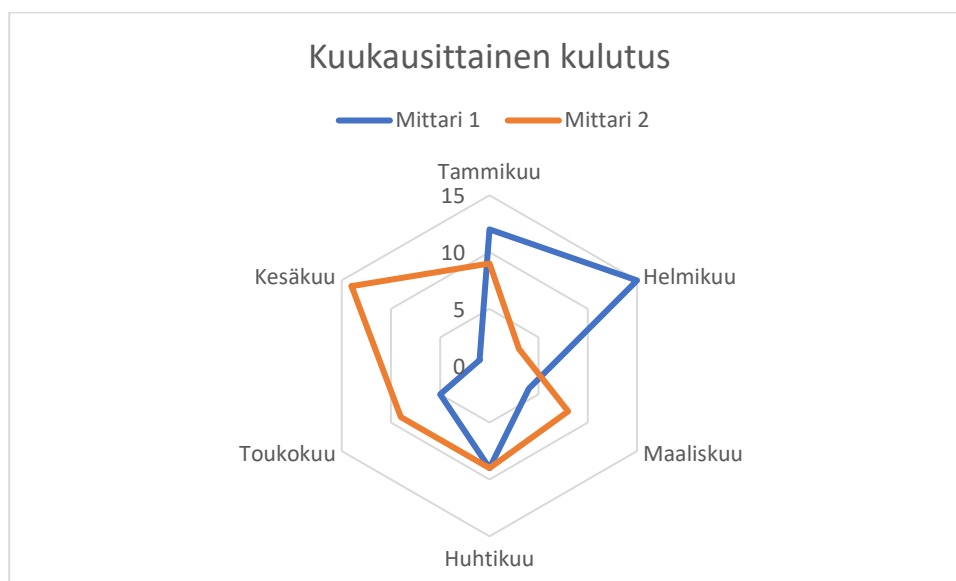
Kuva 11. Ympyräkaavio suhteelliset osuudet (määrät ovat mielikuvituksellisia eivätkä todellisia!).

Kyseinen kaavio on tehty käyttämällä Microsoft Officen Excel-taulukko ohjelmaa. Ohjelma järjestää osuudet automaattisesti suurimmasta pienimpään niin, että suurin osuus alkaa kohdasta kello 12. Tässä ympyräkaaviossa on esitelty mielikuvituksellisia suhteellisia osia kaupunkien välillä. Kaaviosta nähdään, että suhteellinen osuus on suurin Helsingissä ja toisaalta pienin Limingassa. Piirakkakaaviota sovelletaan hyvin paljon ja suurimman osuuden aloitus kello 12 on vain yksi tapa. Eri työkalut järjestävät osuudet eri tavalla, mutta suurimman aloitus kello 12 on yksi yleisimmistä.[30, 32, 36]

2.10.5. Tutkakaavio

Tutkakaaviolla voidaan esittää useampien muuttujien dataa samassa kaksiulotteisessa kuviossa. Kuviossa muuttujia on vähintään kolme, usein kuitenkin enemmän, ja ne

muodostavat kuviosta monikulmion. Jokainen kulma kuvaa omaa muuttujaa ja samaan aikaan kuvion ulottuvuus tiettyihin kulmiin kuvaa suuruutta. Kuvasta 12 voidaan havainnoida kulmat ja niihin sijoitetut eri kategoriat.[37]

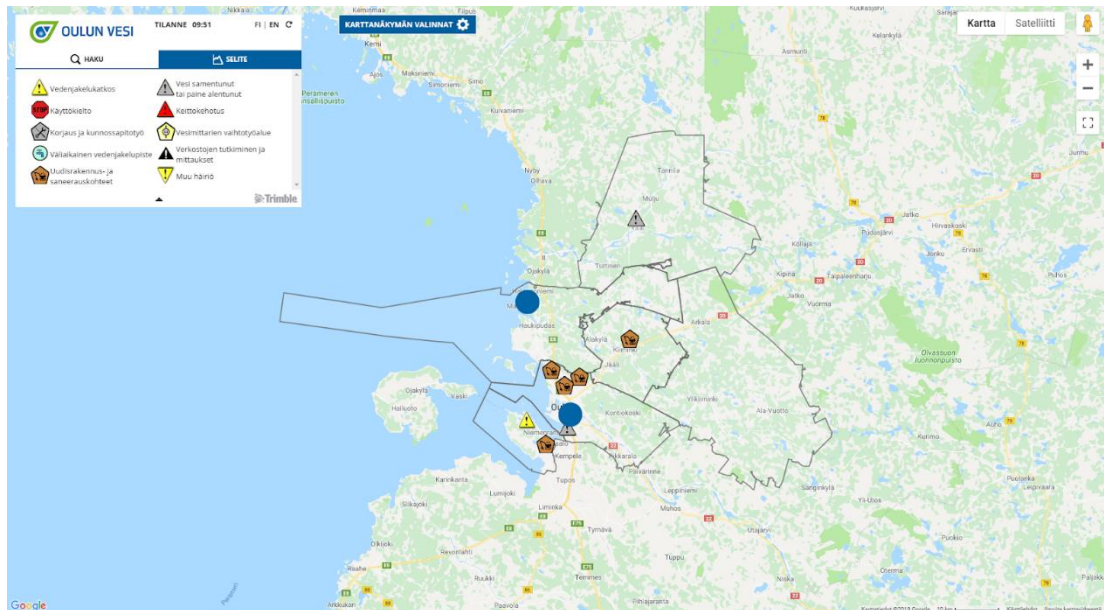


Kuva 12. Tutkakaavio.

Kuvassa 12 kategoriat ovat kuukaudet, jotka löytyvät tutkakaavion laidoista. Asteikko 0-15 kuvaa kulutuksen määrää kyseisenä kuukautena. Kun tarkastellaan kaaviosta sinistä kuviota, huomataan että kulutus tammikuussa on ollut 13m^3 ja helmikuussa 15m^3 . Toisaalta kaaviosta on helppo tulkita samalla missä kuussa vedenkulutus on ollut suurinta ja missä taas pienintä. Tutkakaavio on oivallinen vaihtoehto, kun halutaan nähdä mille osa-alueille määrät ovat asettuneet.[30]

2.10.6. Datakartat

Datakarttaratkaisut ovat vain yksi esimerkki siitä, miten visualisointia voidaan käyttää hyödyksi mallintamaan ja yksinkertaistamaan suuria määriä tietoa kartan avulla. Karttaa käytetään kuvaamaan erilaisia asioita ja sitä on yksinkertaistettu näyttämään vain tarvittava informaatio. Kun käytetään karttaan apuna, voidaan käyttää myös muita menetelmiä hyväksi kuten kuvioiden muotoja. Kolmio, viisikulmio, ympyrä ja niiden värien vaihtelu ja muut variaatiot voivat kuvata erilaisia asioita kartassa. Esimerkiksi Oulun Veden häiriökarttapalvelusta lainatussa kuvasta 13 nähdään, kuinka symboleita on käytetty kuvaamaan eri asioita.



Kuva 13. Trimblen tarjoama häiriökartta palvelu Oulun Vedelle [38].

Kartasta nähdään selvästi häiriösjainti, mutta samalla symboleita ja värejä käyttämällä, minkälainen häiriö on kyseessä. Värit kuvaavat ongelman kriittisyyttä. Punainen näyttää erittäin kriittiseltä, keltainen normaalilta häiriöltä ja harmaa ilmoitukselta.

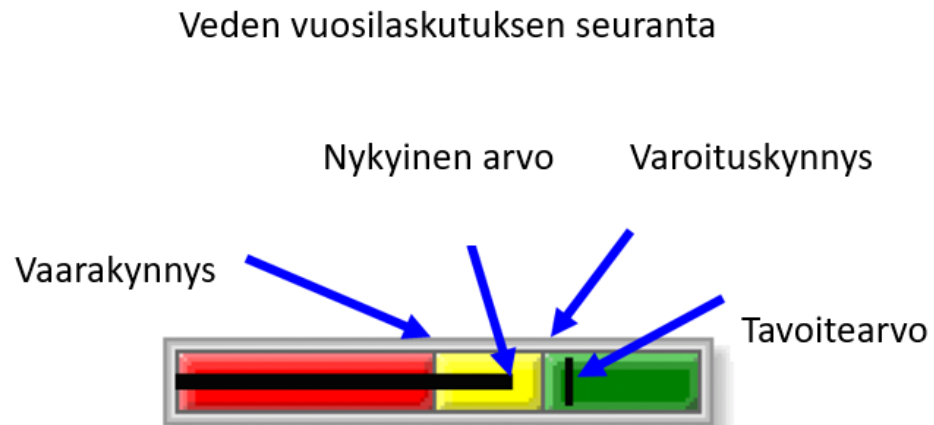


Kuva 14. Häiriökartan symbolit ja kuviot tarkemmin [38].

Toisaalta kuviot kiinnittävät huomiota samalla tavalla kuin liikennemerkkit. STOP-merkki on syytä katsoa, samoin eri variaatiot kolmioista. Toisaalta muut kuviot eivät näytä niin kriittisiltä ja ne voidaan sivuuttaa, ellei haluta tutkia hiukan enemmän (kuva 14).[32]

2.10.7. Bullet-kaavio

Bullet-kaaviolle on yleistä seurata asian edistymistä kaaviota hyväksi käyttäen. Tyypillistä kaaviolle on värien käyttö, jotta erotetaan hyvin, kun jokin tietty etappi (edistymisen kriteeri) on saavutettu (kuva 15).

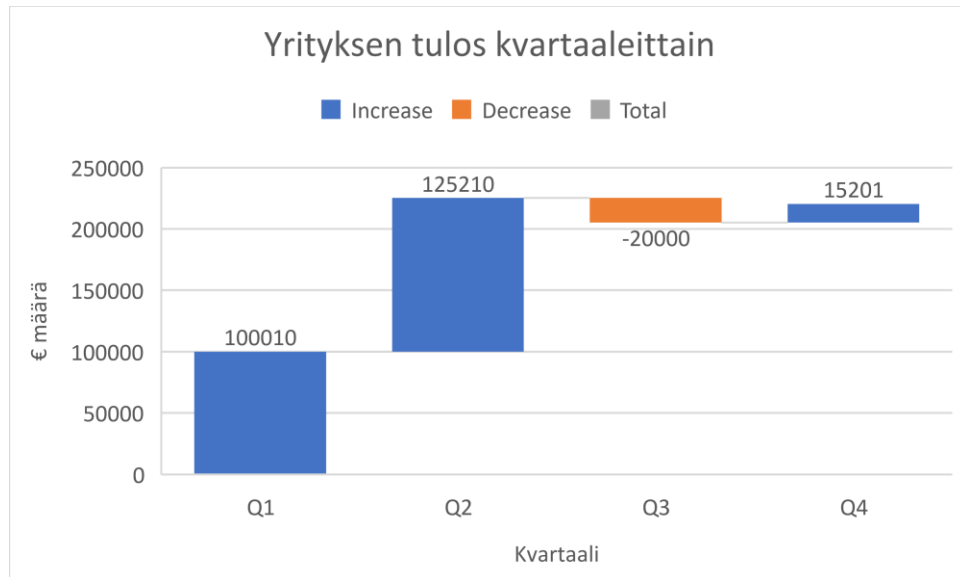


Kuva 15. Bullet-kaavio [39].

Tarkastellessa yrityksille kriittistä osa-aluetta, kuten yrityksen liikevaihdon kehitystä, voidaan käyttää Bullet-kaaviota. Kuvassa 15 seurataan veden vuotuista laskutusta ja siitä syntyvää liikevaihtoa. Nykyistä tilannetta liikevaihdon suhteen kuvaa musta viiva (kuvassa 15, Nykyinen arvo). Kun ollaan punaisella alueella (kuvassa 15, Vaarakynnys), ollaan liikevaihdon suhteen jäljessä ja mahdollisesti tehdään tappiota. Keltaisella alueella ollaan neutraalilla alueella. Tällöin yritys olisi nolatilanteessa, mikä tarkoittaa, että yritys ei tee tappiota eikä voittoa. Kun päästään tavoitteeseen (kuvassa 15, Tavoitearvo), ollaan tilanteessa, johon on pyritty suunnitelmien perusteella. Mustan viivan ylitettyä tavoiteviivan ollaan erittäin hyvissä asemissa ja tehdään suurta liikevaihtoa. Kuvasta on siis helppo seurata ajan funktiota – joskin se vaatii hiukan enemmän tietämystä siitä, kuinka kaavio toimii ja miten sitä hyödynnetään.[30]

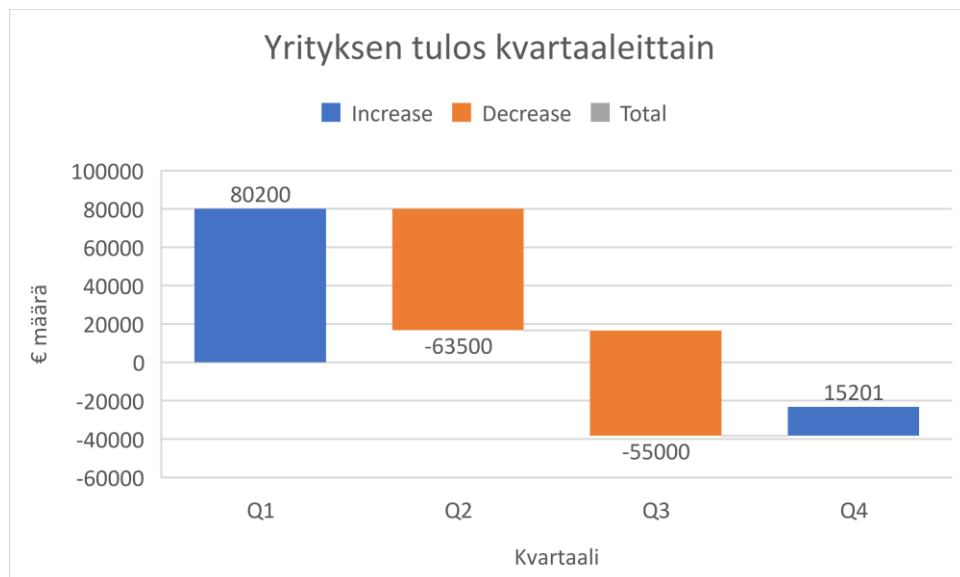
2.10.8. Vesiputous-kaavio

Vesiputous-kaavio on yksi variaatio pylväsdiagrammista. X-akselilla kuvataan usein ajanjaksoa, kun taas y-akselilla voi olla eri arvoja. Kuvassa 16, x-akselilla on kuvattu yrityksen neljännekset (Q1, Q2, Q3, Q4). Y-akselilla on kuvattu tuloksen suuruutta (numeerinen) jokaisella neljänneksellä.



Kuva 16. Vesiputous-kaavio (määrät ovat mielikuvituksellisia eivätkä todellisia!).

Kuvaa 16 selventävät värit, jossa sininen kuvaa positiivista tulosta ja oranssi negatiivista tulosta. Suurin huomioitava ero normaaliin pylväskaavioon verrattuna on palkkien aloituspaikka y-akselilla. Ensimmäinen neljännes alkaa nolasta ja siitä seuraava neljännes aina siihen mihin edeltävällä neljänneksellä on päädytty (harmaa summaviiva, erittäin ohut viiva palkkien välissä). Näin pylväs kasvaa positiivisen tuloksen myötä ja laskee negatiivisen tuloksen myötä. Negatiivisten tulosten ollessa suurempia kuin positiivisten, pylväs voi jäädä myös nollan alapuolelle eli olla negatiivinen kuten kuvassa 17.[40]



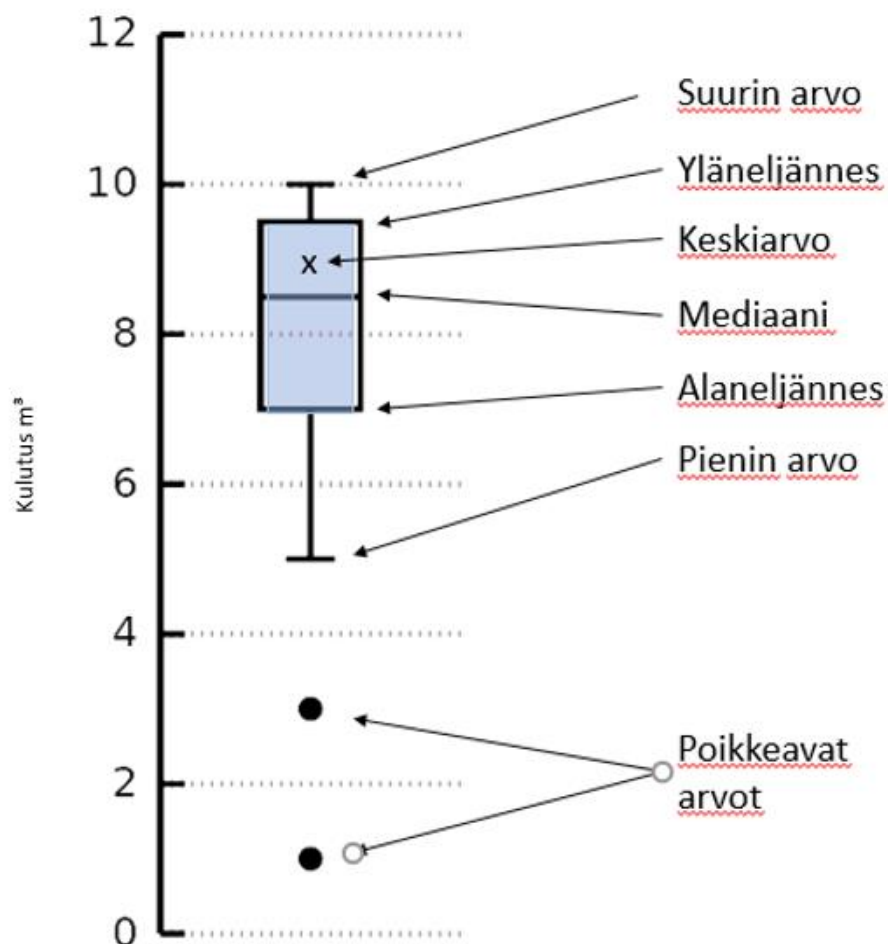
Kuva 17. Vesiputous-kaavio (negatiivinen, määrät ovat mielikuvituksellisia eivätkä todellisia!).

2.10.9. Laatikkokaavio

Laatikkokaavion avulla voidaan näyttää visuaalisesti arvojen jakautumista mitattavasta datajoukosta. Kuviossa alareuna kuvaa alaneljänestä ja yläreuna yläneljänestä. Ruudun sisälle piirretty viiva kuvaa mediaania ja ruudun sisällä oleva rasti keskiarvoa. Viivojen eli janojen ylä- sekä alapää vastaavat suurinta ja pienintä arvoa. Mikäli kuvassa on ulkopuolisia pisteitä, niitä kutsutaan poikkeaviksi arvoiksi. Arvo on poikkeava, mikäli se on yli 1,5 korkeuden päästä laatikon reunasta. Kuvasta 18 nähdään, missä kaikki nämä sijaitsevat.[41]

Laatikkokaavio on vaihtoehto, kun halutaan tutkia jonkin mittarin arvojen jakaumaa. Laatikkokaavioista on erityisesti helppo nähdä, mikäli lukemista löytyy poikkeamia. Poikkeamat ovat yleisesti ottaen virheellisiä lukemia eli lukeman lähettämisen, vastaanottamisen tai käsittelyssä on tapahtunut virhe. Jos näin ei ole, voidaan olettaa, että kohteessa on todennäköisesti tukos tai vuoto, jonka takia poikkeama näkyy selvästi laatikkokaaviosta.

Poikkeamien havaitseminen kulutuksessa



Kuva 18. Laatikkokaavio.

2.11. Kirjastot ohjelmointiin

Ohjelmakirjastojen avulla voidaan luoda edellisessä kappaleessa mainittuja kaavioita. Kirjastot on suunniteltu usein useamman ihmisen toimesta ja niitä ylläpidetään vuodesta toiseen. Kirjastojen käyttö vaatii kuitenkin osaamista tietotekniikan parissa – erityisesti kirjaston sisältävän lisenssin tunteminen on tärkeää, kun tuotetta, jossa kirjastoa käytetään, lähdetään myymään. Eri lisenssit antavat oikeudet kaupallistamiseen eri ehdoin.

2.11.1. *Chart.js*

Chart.js on javascript (js) perusteinen kirjasto, jota voidaan hyödyntää mallintamaan dataa kaavioina. Chart.js on erittäin hyödyllinen työkalu, jolla pystytään tekemään tietystä datajoukosta useita erityyppisiä kaavioita. Chart.js tukee tällä hetkellä (vuonna 2019) yhdeksää erilaista kaaviotyyppiä. Kaavioiden lukeutuu esimerkiksi viivakaavio, pylväskaavio, tutkakaavio, piirakka- ja donitsikaavio, aluekaavio ja yhdistelmäkaavio. Yhdistelmäkaavio osaa näyttää esimerkiksi kahta erilaista kaaviota samassa kaaviossa kuten pylväs- ja viivakaaviota. Kirjasto on kooltaan pieni (puhutaan kilotavuista), joten se vie projektissa vähän levytilaa. Kaiken tämän lisäksi Chart.js on myös responsiivinen. Responsiivisella tarkoitetaan, että kaaviot mukautuvat näytölle riippuen vapaasta tilasta näytöllä. Näytön ja tilan ollessa pienempi kaavio mukautuu siihen automaattisesti ja vastaavasti näytön ollessa isompi kaavio on myös suurempi.[42]

Dokumentaatio, jonka Chart.js tarjoaa nettisivuillaan (<https://www.chartjs.org/docs/latest/>), on erittäin selkeä ja kattaa kaiken tarvittavan informaation mahdollisten kaavioiden käytöstä. Chart.js tarjoaa kattavan esimerkki valikoiman omista kaavioistaan, joihin kehittäjä voi tutustua. Dokumentaatio tarjoaa myös ohjeet, kuinka ja miten data tulee syöttää kaavioon, jotta se näyttää halutun kaavion oikein. Chart.js on oiva vaihtoehto, kun ohjelma kirjastoksi halutaan avoimen lähdekoodin kirjasto, jota hyväksi käyttämällä pystytään luomaan useita erikaavioita. Chart.js on avointa lähdekoodia ja saatavilla MIT-lisenssillä. MIT-lisenssi tarkoittaa, että kyseistä kirjastoa saa käyttää mihin tahansa tarkoitukseen, kuten kaupallistamiseen ilman erillisiä lupia.[43]

2.11.2. *D3.js*

D3.js on web-pohjainen datan visualisointiin perustuva työkalu. D3.js on helppokäyttöinen ja se on mahdollista integroida lähes mihin tahansa web-pohjaiseen ohjelmaan. D3.js on rakennettu javascriptin päälle ja tarjoaa samalla tavalla chart.js:än kanssa kattavan dokumentaation. D3.js:n avulla voidaan luoda nopeasti ja helposti kaavioita suuresta datajoukosta. D3.js on avointa lähdekoodia ja saatavilla BSD-lisenssillä. BSD-lisenssin käyttöoikeudet ovat muuten samat kuin MIT-lisenssillä, mutta lisenssin sisältämä teksti on pidettävä lähdekoodissa mukana.[44]

2.11.3. *Dygraphs*

Dygraphs on avoimen lähdekoodin javascript-kirjasto ja varustettu MIT-lisenssillä. Dygraphsia käyttäen pystytään tutkimaan tiheitä tietojoukkoja. Dygraphs tukee vain

viivakaaviota ja sen se tekee erittäin hyvin ja kattavasti. Dygraphs mainostaa omilla nettisivuillaan, että se pystyy piirtämään miljoonia datapisteitä ilman jumiutumista. Dygraphs onkin oiva työkalu viivakaavioiden tekoon erityisesti silloin, kun halutaan tarkemmin tiedustella suurta datajoukkoa tietyllä aikavälillä. Viivakaavioissa voidaan näyttää useita datajoukkoja samassa kaaviossa ja verrata niitä toisiinsa. Käyttö on helppoa ja sujuvaa, mutta vaatii hieman opettelua.[45]

2.12. Mitattavat laitteet

Aiheen rajauksen myötä keskitytään vain vesihuollossa oleviin mittalaitteisiin. Ennen kaikkea tärkeintä on löytää vesihuollon parissa tarvittavia mittalaitteita, jotka tuovat lisäarvoa vesilaitoksille. Mittalaitteiden tulee olla älykkäitä laitteita, jotta pystytään maksimoimaan ja optimoimaan digitalisoitumisen hyöty. Vesihuollon piirissä ollaan siirtymässä etäluettaviin vesimittareihin. Oulun Energialta saadun tiedon mukaan etäluettaviin mittareihin on siirrytty lämpöenergian puolella Oulussa vuonna 2012, jolloin viimeisimmät mittarit on saatu vaihdettua. Lisäksi yhä useampi vesihuoltolaitos on siirtymässä tai vaihtoehtoisesti testausvaiheessa etäluettavien mittareiden suhteen.

Kuten mainittu, vesihuollossa ollaan tällä hetkellä siirtymävaiheessa ja mahdollisia investointeja etäluettaviin mittareihin mietitään. Mahdollisia investointeja halutaan miettiä rauhassa ja ajan kanssa, jotta vältetään vääriltä päätöksiltä. Osa vesihuoltolaitoksista, kuten Limingan vesihuolto, on jo tehnyt suunnitelmat ja aloittaneet mittarien vaihdon. Osa vesihuollon laitoksista on liikenteessä piloteilla (vain muutama etäluettava mittari), kun taas toiset laitokset ovat tehneet suunnitelmat pitemmälle. Suunnitelmat kattavat erilaisia asioita, kuten missä aikataulussa mittarit vaihdetaan ja mitä järjestelmiä halutaan etäluettaville mittareille, jotta kerätty data voidaan hyödyntää. Tällä hetkellä vesihuollossa vesimittarista luetaan ainoastaan vesimittarin lukema ja tämäkin tehdään manuaalisesti. Mittaus ilmoitetaan yleisesti asiakkaan toimesta kerran vuodessa, mutta poikkeustapauksissa välillä useammin ja välillä harvemmin. Näin ollen arvokasta ja varsinkin reaaliaikaista dataa menee hukkaan. Siirryttäessä etäluettaviin vesimittareihin, on mahdollista kerätä myös muuta kuin pelkästään vedenkulutuslukemia.

2.13. Vastaanotettu data

Kamstrup mainostaa älykästä etäluettavaa mittaria (MULTICAL 21), joka käyttää hyväkseen ultraäänitekniikkaa. Tarkkuus mittarissa on Kamstrupin mukaan huippuluokkaa koko käyttöiän, joka voi yltää 16 vuoteen asti. Kyseinen MULTICAL 21 mittari on suunniteltu asuintaloihin, ja mittarilla on tiedonsiirtovalmius ilman lisälaitteiden tarvetta [46]. Mittari tuottaa dataa kolmesta eri kategoriasta, joita ovat määrä, lämpö ja virheet. Kun puhutaan määrästä, tarkoitetaan vedenkulutusta. Kamstrup mainostaa, että mikäli virtaama on niinkin pientä kuin 2 litraa tunnissa, niin mittarin mittalaite käynnistyy ja tekee mittauksia. Tämä toiminto tekee siis saman kuin henkilö, joka ilmoittaa mittarilukeman. Käyttämällä älykästä etäluettavaa mittaria saadaan kulutus mitattua tiheämpään ja pystytään tekemään tarkempia mittauksia. Riippuen laiteyhdistelmästä, voidaan ilmoittaa myös vähimmäis- ja enimmäisvirtaus.[46]

Kuten mainittu, mittari mittaa myös lämpötilaa. Tarkennettuna mittarin mitatessa se tallentaa ja lähettää myös mitattavan veden lämpötilan eteenpäin. Tämän lisäksi

mitataan myös ympäristön lämpötilaa. Mitattavien suureiden määrä kasvaa ja saadaan lisää dataa, jonka avulla pystytään tekemään ennusteita ja arvioita.

Virheet ovat tärkeä mittalaitteen ilmoittama informaatio. Mittari tukee viittä eri infokoodia, joita ovat vuoto, putkirikko, kuiva, väärä virtaussuunta ja peukalointi. Näiden avulla voidaan tutkia ja estää mahdollisia petoksia, pienentää vedenkulutusta ja tarkistaa mittatietojen luotettavuutta sekä mittalaitteen toimivuutta.[46]

Axioma on toinen mittarivalmistaja, joka on käytössä vesihuollon piirissä. Data, jota mittarit lähettävät, ovat ainakin tällä hetkellä täysin riippuvaisia mittarin valmistajasta. Tämä tarkoittaa, että vaikka Kamstrup lähettää virtauksia, ei välttämättä joku toinen mittarin valmistaja samaa tietoa lähetä lainkaan. Toisaalta vaikka molemmat mittarivalmistajat lähettävät samoja parametrejä se, minkälaisessa formaatissa ne tulevat, vaihtelee erittäin paljon. Osa mittarivalmistajien mittareista lähettää bittijonon ja osa taas JSON-tiedoston ja niin edelleen.

2.13.1. Raakadata ja tietämys

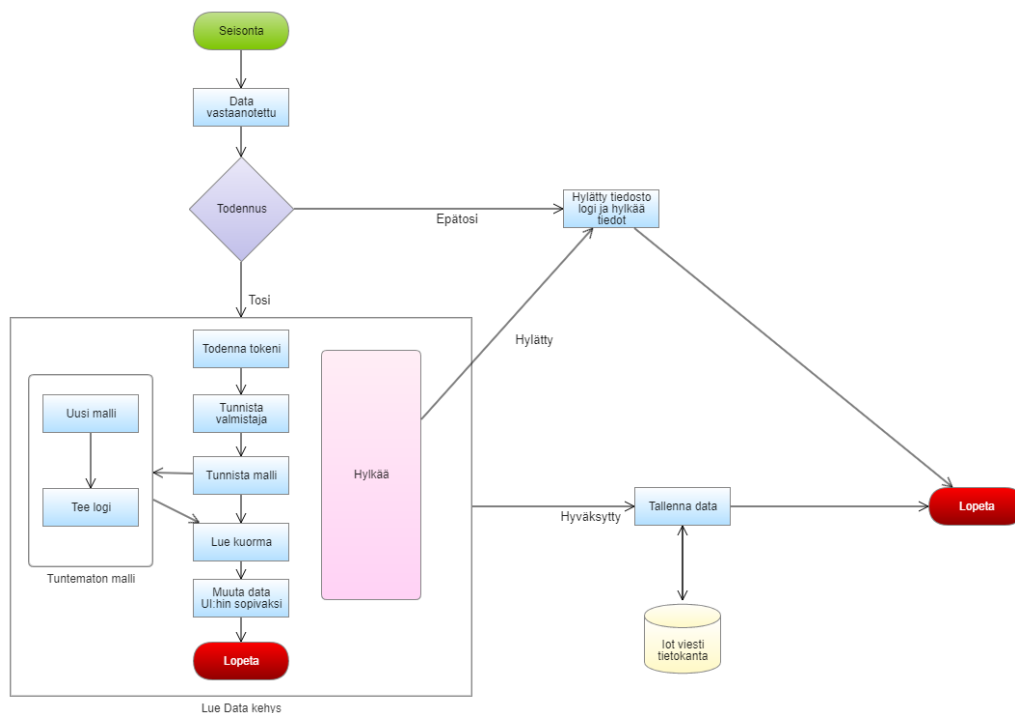
Raw data eli raakadata on dataa, joka tulee mitattavasta laitteesta. Kerättyä dataa kutsutaan raakadatakse silloin, kun sille ei ole vielä tehty minkäänlaisia toimenpiteitä. Tämä tarkoittaa esimerkiksi sitä, että dataa ei ole käsitelty, muokattu tai vaikkapa puhdistettu. Kyseessä on siis vain esimerkiksi päivämäärä, joka on lähetetty laitteesta. Päivämäärä on esimerkkinä erinomainen, sillä ohjelmoinnissa päivämäärä voidaan ilmoittaa monella eri tavalla. Kun päivämäärä otetaan vastaan mittalaitteesta, se voi tulla muodossa 25/1/2019, 25.1.20019, 25 Jan, 25st January 2019 tai vaikkapa 1548374400000 (millisekuntia). Osa tekniikoista on tarkempia kuin toiset, mutta järjestelmästä, koodikielestä ja käyttötarkoituksesta riippuen data pitää aina muokata käyttäjälle ymmärrettävään muotoon. Mittalaitteet voivat lähettää saman parametrin mittalaitteesta ja valmistajasta riippuen täysin eri muodossa. Datalle täytyy tehdä muutoksia ja muokkauksia, jotta se voidaan tallentaa ja muokata sopivaksi järjestelmälle. Toimenpiteen jälkeen ei puhuta enää vain datasta vaan usein informaatiosta. Raakadata on dataa, jonka ymmärtää vain asiasta tietävä, ja informaatioksi kutsutaan dataa, joka on muutettu muotoon, joka antaa datalle kontekstin. Itse raakadata on harvoin hyödyllistä sellaisenaan, mutta sen kerääminen on kuitenkin tärkeää, jotta voidaan muodostaa informaatiota, jota puolestaan voidaan hyödyntää eteenpäin.[47]

Informaation perusteella luodaan tietoa, jota pystytään edelleen jalostamaan tietämykseksi. Kun lähdetään tutkimaan käsitteitä tieto ja tietämys, tulee ensimmäisenä mieleen ihminen. Tämä on myös tärkeä tietää, sillä tieto ja tietämys ovat käytännössä ihmisen ymmärryksen tuotetta. Ihminen on se, joka muodostaa informaatiosta tietoa ja tietämystä, ja tällä hetkellä sen kehitys on ihmisestä kiinni – ei koneista ja niiden algoritmeista.[47]

2.13.2. Rakenne

Datapalvelu keskittyy Digitan tarjoamaan dataan. Digitan lähettämään dataan on rakennettu bittijonon käsittelyyn tarvittavia funktioita valmiiksi Vesitiedon asiakas- ja laskutusjärjestelmässä. Ratkaisuihin on otettu huomioon myös eri valmistajat ja se, miten bittijonon luku sekä vienti tietokantaan informatiiviseen muotoon hoituu.

Bittijonosta luetaan tunnisteista kuten mistä data on peräisin, kenen valmistama mittari on, milloin data on lähetetty, mittarin lukemat ja muut arvot. Kuvasta 19 nähdään datan vastaanotto, tunnistaminen, tallentaminen tai sivuuttaminen Datapalvelussa, mikäli dataa ei voida avata tai tunnistaa.



Kuva 19. Vastaanotetun datan vuokaavio [48].

Datan näyttämistä, analysointia ja visualisointia varten on tärkeää miettiä rakenne oikein tietokantaan. On otettava huomioon, mitä tietoja tallennetaan ja mihin. Täytyy muistaa paitsi optimoida tietokannasta haku, mutta ennen kaikkea tulevaisuudessa laajentuminen ja palvelun tarjoaminen useammalle eri taholle.

Datan analysointia varten tarvitaan oikean tahoon kohdistettu data, kuten esimerkiksi vesihuoltolaitokseen. Tämän lisäksi usein halutaan kohdistaa haku vain tiettyyn asiaan, kuten vesimittareihin. Tämän kaltaista jaottelua hyväksi käyttäen saadaan haettua tietokannasta vain tarvittava data helposti ja analysoinnin tekeminen on järjestelmälle kevyempää. Tämän takia rakenne ja sijoittelu datalle tietokannassa on tärkeää, jotta saadaan tarkasteluun vain tarvittava data.

2.14. Massadata

Big data eli massadata on käsite, jota käytetään suurille määrille dataa. Puhuttaessa datasta tarkoitetaan kaikkea materiaalia, joka on kirjallista, digitaalista sekä visuaalista. Vuosien saatossa on tapahtunut iso muutos datan tallentamisessa. Mitä kauemmas historiassa katsotaan, sitä vähemmän dataa on saatavilla. Teknologia on kehittynyt huimaa vauhtia ja tänä päivänä on mahdollista sekä hallita että tallentaa suuria datamääriä digitaaliseen muotoon turvallisesti. Tallentamisen muotoa on myös

pystytty tuomaan teknologian avulla kaikille mahdolliseksi. Tämä tarkoittaa datamäärien nopeaa kasvamista, johon tarvitaan taas uusia tapoja ja menetelmiä hallitsemiseen. Datamäärien kasvaessa kiihtyvästi ja nopeasti on jossain välissä syntynyt käsite Big Data eli massadata. Tarkkaa ajankohtaa milloin käsite massadata on syntynyt ei ole tiedossa.[47]

Massadatalla tarkoitetaan datamäärää, jota ei pystytä hallinnoimaan yleisesti käytetyillä ohjelmilla. Koska massadata on erittäin heikosti määritelty käsite, se muuttuu koko ajan. Massadatan käsitettä ei siis voida määritellä suoraan datamäärästä. Teknologioden kehittyessä myös massadatan määritelmä elää, riippuen siitä, miten ohjelmistot ja menetelmät kehittyvät.[49]

Kaplan ja Haenlein määrittelevät massadatan kolmen V:n avulla. Volyymi (volume) eli datan määrä, joka tallennetaan, vaihtelevuus (variety) eli datan tyyppi, joka vaihtelee, nopeus (velocity) eli nopeus, jolla data pystytään prosessoimaan, sillä usein puhutaan lähes reaaliaikaisesta datasta. Massadatasta puhuttaessa tarkoitetaan kaikkea mahdollista dataa, jota kerätään, ei pelkästään mittalaitteista kerättyä dataa. Data voi tulla esimerkiksi mistä tahansa laitteista tai ohjelmistoista.[52] Jotkin tahot ovat ottaneet myös neljännen V:n käyttöön, joka kuvastaa datan laatua ja arvoa (Veracity).[51]

3. VESIHUOLTOLAITOKSET

Tässä luvussa käsitellään Suomen vesihuoltolaitosten tilaa tällä hetkellä ja niiden vertailua muihin energia-alan yhtiöihin, kuten Oulun Energiaan. Tarkoitus on tarjota lukijalle pohja vesihuoltolaitosten nykyisestä tilanteesta ja samalla valottaa vesihuollon tulevaisuuden näkymiä. Käsitellään syitä, minkä takia digitalisaatio vesihuollossa on muita energia-aloja, kuten sähkö- ja lämpöenergia-aloja, jäljessä. Keskitymme myös nykyisiin vesihuollon vaatimuksiin ja mitä vesihuollon toivotaan tulevaisuudessa saavuttavan digitalisoitumisen myötä.

3.1. Vesihuoltolaitokset vuonna 2019

Tarkastellessa vesihuoltolaitoksia vuonna 2019 huomataan niiden pystyvän hoitamaan ydinliiketoiminta nykyisillä menetelmillä. Vettä tulee lähes poikkeuksetta jokaiseen talouteen, eikä vedestä ole pulaa samalla tavalla kuin esimerkiksi sähköstä. Samaan aikaa vesihuollon liiketoiminnan kehitystä hidastaa kilpailun puute. Suomessa vesihuoltolaitokset toimivat julkisella sektorilla, mikä tarkoittaa, ettei yrityksen ydinliiketoimintaan kuulu maksimaalisen voiton tavoittelemisen, vaan yksinkertaisesti infrastruktuurin ylläpitäminen eli perusrakenteen yleinen kunnossapito. Uusien järjestelmien rakentamisella tai tarvittaviin investointeihin käyttämiselle ei ole nähty tarpeeksi hyötyjä. Syynä tähän näyttää olevan, että vanhat järjestelmät ja menetelmät ovat olleet riittäviä tuottamaan halutun tuloksen vuodesta toiseen.[53]

Erinäköisten Excel-taulukoiden käyttäminen ja paperilaput ovat arkipäivää vielä monelle vesihuoltolaitokselle. Vesilaskutuksessa käytetään usein arvioperusteista laskutusta, jossa asiakas maksaa vedestä kuukausimaksua taloudessa olevan henkilömäärän mukaan. Tätä laskutusta tasataan vuosittain, joko asiakkaan hyväksi, jolloin asiakkaalle maksetaan rahaa takaisin eli arviolaskutus on ollut suurempaa kuin vedenkulutus. Mikäli asiakas on kuluttanut enemmän vettä vuoden aikana kuin arviolaskutuksen mukaan on laskutettu asiakkaalta, asiakas saa laskun, joka sisältää ylimääräisestä vedestä koostuvan summan. Kuten mainittu aikaisemmin, laskutustapa toimii, mutta on työläs, monimutkainen ja aiheuttaa turhia kuluja vesihuoltolaitoksille. Teknologian kehitys vuosien saatossa mahdollistaisi huomattavasti reaaliaikaisemman kulutuksen seurannan ja sitä hyödyntämällä uuden laskutustavan. Oulun Veden mukaan tämä on yksi syistä miksi vesihuoltolaitokset ovat siirtymässä etäluettaviin vesimittareihin. Vertailun vuoksi vesihuoltolaitokset ovat vasta siirtymävaiheessa etäluettavien mittareiden asennuksessa ja käyttöönotossa, kun taas esimerkiksi Oulun Energiolla kaikki mittarit on vaihdettu lähes poikkeuksetta etäluettaviin mittareihin vuonna 2012.

3.2. Vesihuollon säännökset

Vesihuollon toimiessa julkisella sektorilla sitä säännöstelee valtio. Valtio säännöstelee, mitä vesihuoltolaitosten on toteutettava ja miten. Vesihuoltolaki määrittelee seuraavat asiat:

- Yleiset säännökset
- Vesihuollon kehittäminen ja järjestäminen

- Liittäminen vesihuoltolaitoksen verkostoon ja vesihuollon hoitaminen
- Huleveden viemäroinnin järjestäminen ja hoitaminen
- Vesihuoltolaitoksen taloushallinto
- Sopimukset vesihuollossa ja huleveden viemäroinnistä
- Keskeytys ja virhe
- Valvonta, hallintopakko ja muutoksenhaku
- Erinäiset säännökset
- Muutossäädösten voimaantulo ja soveltaminen

Säännösten ydinidea on pitää huoli vedenjakelusta ja jäteveden oikeanlaisesta käsittelystä ympäristön kannalta. Säännökset takaavat jokaiselle kuluttajalle helposti saatavilla olevan puhtaan juomakelpoisen hanaveden, jossa tuotteen hinta on järkevästi hinnoiteltu. Säännökset pitävät huolen, että veden laadusta pidetään huolta lain vaatimien kriteerien mukaisesti, ja että jätevesi käsitellään oikeanmukaisesti.[52]

3.3. Vesihuollon tarpeet tulevaisuudessa

Digitalisoitumisen myötä vesihuoltolaitosten tarpeet sekä vaatimukset nousevat vesihuollon omasta tarpeesta, eivätkä ainoastaan valtion säännösten muuttumisesta. Ikivanhat tavat pyritään uusimaan portaittain vastaamaan kuluttajien tarpeita entistä paremmin. Sähkö- ja lämpöenergian puolella etäluettavista mittareista saadut hyödyt on nähty ja vesihuoltolaitokset ovat heränneet siihen, että on muutoksen aika. Älykkäät vesimittarit ovat vain yksi askeleista kohti digitalisoitumista vesihuollon parissa. Mittareiden vaihto ei kuitenkaan ole suoraviivaista vaan vaatii paljon suunnitelmia ja investointeja. Esimerkiksi Oulun vedellä on yli 36 000 käyttöpaikkaa, joka tarkoittaa samalla myös yli 36 000 vesimittaria.

Muutos tapahtuu pienissä erissä. Vesilaitokset pyrkivät vaihtamaan tietyn määrän mittareita etäluettaviin mittareihin vuosittain. Määrä vaihtelee vesihuoltolaitoskohtaisesti riippuen investoinneista ja työmäärästä. Mittareiden vaihto aloitetaan pääsääntöisesti vanhimmista mittareista, joilla mittarin käyttöikä on tullut elinkaarensa päähän. Älykäs vesimittari on kalliimpi kuin nykyinen mittari. Etäyhteyden ylläpito ja siihen tarvittava verkkoyhteys ja lukujen lähettäminen päätelaitteelta maksavat. Lisäksi tarvitaan kaikki uudet järjestelmät, jotka tukevat kyseisiä uusia teknologiota. Oulun Vesi vaihtaa keskimäärin noin 3500 vesimittaria vuodessa, vesimittareita ollessa noin 36 000, vaihtoaika nykyisellä vauhdilla on reilu 10 vuotta. Älykkäisiin mittareihin siirtyminen on siis pitkän aikavälin suunnitelma. Etäluettavia mittareita halutaan päästä hyödyntämään heti alusta lähtien, vaikka mittareita olisi vaihdettu vasta muutamia. Tällä tavoin mittareiden vaihto etäluettaviin on helpommin ja hyvin perusteltu. Samalla kustannustehokkaita ratkaisuja voidaan aloittaa suunnittelemaan aikaisessa vaiheessa.

Vesihuoltolaitoksilla on tarve saada työkaluja etäluettavien mittareiden datan seurantaan, tarkasteluun ja analysointiin. Tällä tavoin vesihuoltolaitokset saavat digitalisoitumisen hyödyt käyttöön heti ensimmäisistä asennetuista älykkäistä vesimittareista. Tähän kun lisätään vielä kolmannet osapuolet, kuten Vesitietojärjestelmä ja sen hyödyt, työkalujen määrä kasvaa entisestään. Laskutus- ja asiakastietojärjestelmä tarjoaa vesihuoltolaitoksille palvelun, jonka kautta vesihuoltolaitos voi tarkastella koko yritystoimintaa entistäkin laajemmin. Vääjäämättä datan digitalisoitumisen myötä paperiset kaavakkeet sopimuksista tai

kirjemuodossa lähetetyt kulutuslukemat jäävät pikkuhiljaa historiaan ja muuttavat muotonsa sähköiseksi. Vesihuoltolaitokset pääsevät tarkastelemaan digitalisoitumisen myötä omaa – mutta myös asiakkaiden toimintaa – entistä laajemmin.

4. DATAPALVELU TOTEUTUS

Datapalvelun kehitys alkoi keväällä 2019. Datapalvelun rakentaminen oli ollut suunnitelmissa jo ennen vuotta 2018. Muut projektit ja niiden kiireet sekä kiinnostus Datapalvelua kohtaan eivät kohdanneet tarpeeksi hyvin ennen kevättä 2019.

Datapalvelua on lähdetty kehittämään, sillä vesihuollon puolella on nähty tarpeita kyseiselle järjestelmälle. Datapalvelun tuomat ratkaisut on suunniteltu IoT-laitteille ja reaaliaikaisen datan tarkasteluun. Potentiaalisten asiakkaiden mielipiteitä ja tarpeita on kuunneltu läpi projektin, jotta Datapalvelu vastaa mahdollisimman hyvin vesihuollon tulevia ja nykyisiä tarpeita. Datapalvelu on kehitysvaiheessa ja ensimmäisiä demoja on pystytty näyttämään potentiaalisille asiakkaille ennen joulua 2019. Toteutuksessa on käytetty hyväksi erittäin paljon datan visualisointia ja sen variaatioita.

4.1. Yhteistyö

Eri kokousten ja hankkeiden yhteydessä vesihuoltolaitokset ovat kertoneet tarpeitaan liittyen vesihuoltoon. Yksi näistä hankkeista on AIF Water Ecosystem. Kyseisessä hankkeessa Oulun Vesi on tuonut esille omia tarpeitaan, mutta painottanut myös, että pienemmillä vesihuoltolaitoksille on mahdollisesti erilaisia tarpeita ja heidät tulisi ottaa huomioon, kun ratkaisuja suunnitellaan. Tiiviin yhteistyön ansiosta Vesitieto on käynyt Datapalvelua ja sen kehitystyötä jonkin verran läpi Oulun Veden kanssa. Vesitieto on kiinnostunut siitä, minkälaisiin haasteisiin Datapalvelun tulisi tarjota ratkaisuja. Samalla on tiedusteltu, miltä ensimmäiset versiot Datapalvelusta näyttäivät potentiaalisten asiakkaiden silmissä. Datapalvelun ensimmäisiä demoversioita on näytetty Lempäälän vesihuollolle, Limingan vesihuollolle sekä jo mainitulle Oulun Vedelle. Kaikissa vesihuoltolaitoksilla vastaanotto on ollut myönteinen ja Datapalvelulle on nähty suurta potentiaalia.[53]

Oulun veden kanssa käytyjen keskusteluja hyödyntäen Vesitieto on pystynyt poimimaan tärkeitä kriteereitä ratkaisujen kannalta. Datapalvelua on kehitetty yritysten tarpeiden mukaan ja kehitystyötä on tehty esimerkiksi Oulun Vedellä käydyn haastatteluun perustuen. Haastattelu suoritettiin kokoushuoneessa häiriötekijöiden minimoimiseksi. Haastattelua varten oli laadittu kysymyslista, jonka laatimiseen haluttiin useampi henkilö mukaan, sillä osalla henkilöistä on enemmän taustatietoa vesihuollosta ja Oulun Veden toiminnasta kuin toisilla. Haastatteluun osallistujille (Tero Kilpeläinen, Sara Alanära) annettiin mahdollisuus tutustua kysymyksiin etukäteen.

Haastattelussa käytettiin hyväksi puoliksi-jäsenneltyä haastattelutyyppeä, kuitenkin sitä tarpeen vaatiessa soveltaen. Puoliksi-jäsennelty haastattelutyyppeä valittiin, sillä haluttiin kerätä aiheesta syvällisempää tietoa. Haastattelun aikana tiedusteltiin myös muita asioita, joten puoliksi-jäsennelty haastattelutyyppeä tuntui parhaalta valinnalta. Kysymyslista käytiin pääsääntöisesti läpi ja tarvittavat ongelmat saatiin poimittua haastattelun yhteydessä. Haastattelussa haluttiin tiedustella, mitä Oulun Vesi odottaa Datapalvelulta. Vesitieto oli myös kiinnostunut siitä, minkälaisia palveluita palvelun tulisi tarjota, jotta Oulun Vesi haluaisi ostaa palvelun tulevaisuudessa. Samalla haastattelussa tiedusteltiin, miltä Datapalvelun tarjoamat palvelut näyttäivät ostajan silmissä.

Alueellinen kulutus ja sen seurannan puuttuminen nousi suurimmaksi ongelmaksi, johon kaivataan ratkaisuja. Palaute oli positiivista niin ulkoasun kuin nykyisten

toimintojen osalta, joskin joitain parannuksia ja lisäominaisuuksia ehdotettiin. Datapalvelua on esitetty myös jälleenmyyjille, kuten Onniselle ja samankaltaisia tarpeita on noussut esille myös heidän kanssaan käydyissä keskusteluissa.[53]

Oulun Vedellä ei ole tällä hetkellä (joulukuu, 2019) tarkkaa aikataulua, budjettia tai suunnitelmia siitä, miten ja kuinka nopeasti vesimittarit tullaan vaihtamaan etäluettaviin vesimittareihin. Oulun Vedellä on noin 80 etäluettavaa vesimittaria pilotissa, jossa he kokeilevat etäluettavia mittareita ja LoRa verkon toimintaa. Päätökset siitä, mihin verkkoteknologiaan ja mihin etäluettaviin mittareihin päädytään, tehdään myöhemmin ja harkiten, sillä investointi on suuri. Mainittakoon myös, että mikäli mittari, joka päädytään valitsemaan, on väärä ja ei tuota haluttua lopputulosta, mittareihin kohdistuva uudelleeninvestointi on erittäin suuri. Näiden seikkojen valossa johtopäätöksiä ei vedetä hätiköidysti vaan harkiten.[54, 55]

4.2. Ensisijaiset tarpeet

Vesihuoltoa halutaan kehittää eteenpäin kustannustehokkaasti. Tämä tarkoittaa yksinkertaistettuna sitä, että uusien ratkaisujen käyttöönoton täytyy tuoda taloon enemmän rahaa, kuin ratkaisuun investoidaan. Samaan aikaan painotetaan, että ratkaisujen tulisi pystyä hyödyntämään jo olemassa olevia laitteita sekä järjestelmiä mahdollisimman tehokkaasti. Oulun Vesi painottaa olemassa olevien järjestelmien tuottaman datan hyödyntämistä.

Sana integraatio on ollut esillä useaan otteeseen, eikä vanhoja toimivia järjestelmiä haluta uusia. Laitteisiin ja eri järjestelmiin on investoitu suuria summia rahaa ja niitä halutaan hyödyntää myös muualla, kuten Datapalvelussa. Kiteytettynä, vahvimpana valttina Datapalvelulle nähdään jo olemassa olevien järjestelmiä ja laitteiden tarjoaman datan hyödyntäminen eri rajapintojen kautta. Samaa asiaa on painotettu AIF Water Ecosystem hankkeessa. Vesihuoltolaitokset eivät halua lukemattomia uusia järjestelmiä vaan yhden järjestelmän, joka voisi hyödyntää jo olemassa olevia järjestelmiä ja niissä olevaa tietoa.[54, 55]

Datapalvelulle toivottaisiin myös Oulun Veden toimesta jälleenmyyntiarvoa kompensoimaan alkuinvestointia, mutta myös parantamaan kustannustehokkuutta. Tällä tarkoitetaan jonkin osan myymistä Datapalvelusta suoraan kolmannelle osapuolelle, olettaen sen olevan mahdollista. Esimerkkinä Oulun Vesi olisi mahdollisesti kiinnostunut myymään kulutustietojen seurantaan isännöintipalveluille. Tämän mahdollistamiseksi lainsäädäntöön liittyviä kysymyksiä ja ratkaisuja on osittain alettu selvittämään, mutta toimenpiteet vievät kauan aikaa byrokratian takia. Kiteytettynä, Oulun veden ideologiaan kuuluu palveluiden myynti Datapalvelua hyödyntäen myös muille tahoille lain sallimissa rajoissa.[54, 55]

Datapalvelussa jo oleviin ratkaisuihin on tullut parannusehdotuksia sekä uusia innovaatioita. Ideat ovat peräisin Oulun Vedeltä, Onniselta, Limingan Vesihuollolta, Lempäälän Vesihuollolta, että Digitalta. Datapalvelun tuomaan hälytysratkaisuun toivotaan parempaa priorisointia. Hälytyksiin toivotaan värien käyttöä, kuten punaista mikäli hälytyksen prioriteetti on kriittinen. Hälytykset halutaan järjestykseen kriittisyyden mukaan (Datapalvelu mahdollistaa tämä tällä hetkellä) [54, 55]. Perusteluina käytettiin, että ensimmäiset näkyvät hälytykset ovat ne mihin kiinnitetään eniten huomiota ja kriittiset hälytykset halutaan hoitaa nopeasti suurempien vahinkojen välttämiseksi.[54, 55]

Karttapalveluun on toivottu sekä esitelty monia uusia ratkaisuja. Tärkeimpänä ominaisuutena Vesitiedon keräämän tiedon mukaan on alueellinen seuraaminen

vesimittarien osalta. Tähän lukeutuu monia asioita kuten vedenkulutus, vuotovesien määrä ja korkeustasojen vaihteluja. Oulun Vesi on nostanut vuotoveden määrän seurannan korkealla, sillä tätä vettä ei voida laskuttaa asiakkaalta. Tero Kilpeläinen Oulun Vedeltä arvioi vuotoveden määrän olevan 10-14% luokkaa. Pelkästään vuotoveden määrän pienentämisellä ja seuraamisella Oulun vedellä on mahdollisuus kasvattaa liikevaihtoa ja siirtää vesi hyötykäyttöön eikä suoraan jätevedeksi.

4.2.1. Toteutusprioriteetit

Integraatioiden tekeminen eri järjestelmiin ja rajapintojen rakentaminen vaatii paljon yhteistyötä monien eri osapuolten kanssa. Integraatioissa priorisoidaan sellaisia järjestelmiä, jotka hyödyntävät ensisijaisesti suurta määrää potentiaalisista asiakkaista. Integraatioratkaisuja ei lähdetä alustavasti rakentamaan vain yhden asiakkaan tarpeisiin, sillä hyötysuhde on yksinkertaisesti liian pieni. Projektiin vaikuttaa myös se, että yhtään varmaa asiakasta Datapalvelulla ei tässä vaiheessa ole. Integraatioita muihin järjestelmiin tullaan hyvin todennäköisesti toteuttamaan, mutta vasta tulevaisuudessa.

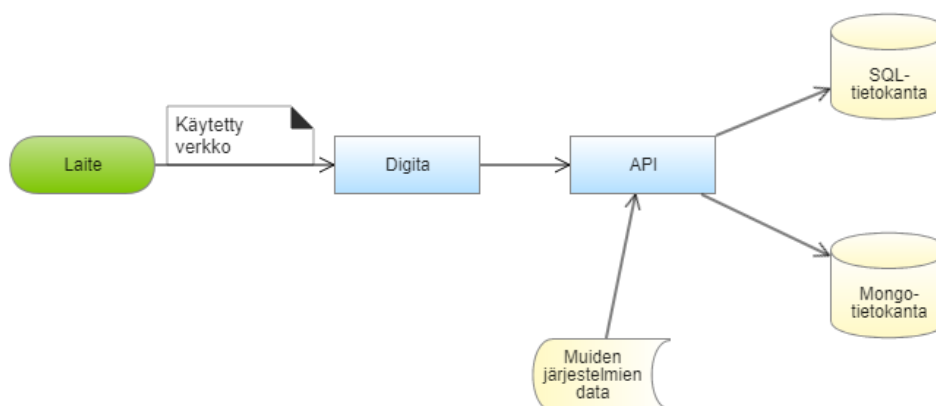
Toteutusjärjestyksen prioriteeteissa korkeimmalle nousee alueellinen seuranta ja hälytysten priorisointi ja niiden seuranta. Toimintoihin lähdetään tekemään parannuksia ja uusia käyttöliittymiä lisää. Tällä tavoin Datapalvelu tuottaa kannattavan ja kustannustehokkaan palvelun vesihuollolle. Yhteistyö eri vesihuoltolaitosten ja muiden osapuolten kanssa jatkuu ja sopivia ratkaisuja ongelmiin pyritään löytämään yhdessä ajan kanssa.[2]

4.3. Määrittely

Datapalvelun tulee olla itsenäinen järjestelmä, joka kuitenkin pystyy hyödyntämään kolmansien osapuolten järjestelmiä. Itsenäisellä järjestelmällä tarkoitetaan järjestelmää, joka on käytettävissä riippumatta muista Vesitiedon kehittämistä ympäristöistä tai järjestelmistä, kuten Vesitiedon laskutus- ja asiakastietojärjestelmästä. Datapalvelu toimii itsenäisesti, vaikka laskutus- ja asiakastietojärjestelmän ylläpito loppuisi tai järjestelmä kaatuisi syystä tai toisesta.

Datapalvelu on suunniteltu vesihuollon pariin, sillä Vesitiedon asiakkaisiin kuuluu useita vesihuoltolaitoksia. Datapalvelun suunnittelussa on otettu huomioon mahdollinen laajentuminen energiasektorille (lämpöenergia ja sähköenergia). Tämän työn ohessa pysytään kuitenkin vesihuollon parissa ja mahdolliset laajentumiset toteutetaan myöhemmin tulevaisuudessa.

Datapalvelun tarkoitus on kerätä ja tallentaa dataa etäluettavista sensoreista. Ensin keskitytään vesimittareihin ja hiljalleen siirrytään kohti muita sensoreita, kuten painesensoreita ja lämpötilasensoreita. Vesimittareilla aloittaminen on luontevaa, sillä Vesitiedolla on rakennettu etäluettavista vesimittareista dataa keräävä järjestelmä jo aikaisemmin. Hyödyntäen kyseistä tietoa, on mahdollista nopeuttaa Datapalvelun käyttöönottoa sekä kehitysvaihetta. Tällä hetkellä laitteen eli mittarin ja Datapalvelun välissä on Digita ja verkkoteknologia. Mikäli verkkoteknologia on luotettava, kuten LoRa ja Digitan päässä ei tapahdu ihmeellisyyksiä, Datapalvelun tarjoamaan API:iin tulee eheää dataa, joka voidaan tallentaa tietokantoihin (Kuva 20) jatkotoimenpiteitä varten. API pystyy tulevaisuudessa keräämään dataa myös muista tietolähteistä.



Kuva 20. Vuokaavio laitteesta tietokantaan.

Kerätty data säilötään eli tallennetaan Azure pilvipalveluun eri tietokantoihin riippuen datasta ja siitä, mitä se pitää sisällään. Eri tietokantaratkaisuina käytetään SQL ja Mongo ratkaisuja, kuten kuvassa 20. Kahteen eri tietokantaratkaisuun on päädytty, koska osa datasta on staattista ja osa dynaamista. Azure pilvipalvelu tarjoaa lähes rajattomat mahdollisuudet säilötyn datan määrälle, mutta myös molemmilla tietokantaratkaisulle.

4.3.1. Tietokanta

Tämä työ ei keskity tietokantaratkaisuihin, mutta Datapalvelun ymmärtämisen vuoksi työ esittää käytetyt ratkaisut. SQL-tietokantaa käytetään staattisen datan tallentamiseen. Staattisella datalla tarkoitetaan esimerkiksi käyttäjiä tai käyttäjäryhmiä. Nämä tietokantarakenteet tulee pysymään hyvin samanlaisena koko Datapalvelun käyttöiän ja vaativat todennäköisesti vain vähän muutoksia. Tilanne on eri, kun puhutaan IoT-mittareissa, sillä data ja sen parametrit muuttuvat valmistajasta ja toisaalta myös palveluntarjoajasta, kuten Digitasta riippuen.

Tietokantarakenteiden ollessa dynaamisia, kuten IoT-mittareissa lähetetyssä datassa, käytetään Mongo tietokantaratkaisua. Mongo on hyvä ratkaisu, sillä se ei välitä datasta tai sen muodosta. Näin ollen uusien parametrien saapuessa ne voidaan tallentaa tietokantaan ja niitä voidaan käsitellä myöhemmin, mikäli sille nähdään tarvetta. Käytössä on Microsoft Azure pilvipalvelun tarjoama Azure Cosmos DB. Kaiken datan pitäminen samassa järjestelmässä, eli Microsoftin tarjoamassa Azure pilvipalvelussa, tuntuu luontevalta ratkaisulta, sillä palvelu tukee Vesitiedon molempia valittuja tietokantoja, mutta myös koko projektin säilömistä Azure pilvipalvelussa.

4.3.2. API

Datapalvelun on pystyttävä vastaanottamaan dataa monesta eri lähteestä. Mitä yhteyteen tulee, sen täytyy olla turvallinen ja vakaa. API:n tehtävänä on pitää huoli

etenkin metodeista GET, PUT, POST, DELETE. Turvallisuudella tarkoitetaan, että metodien käyttö on sallittu vain niille tahoille, joille oikeudet on myönnetty. Vakaalla yhteydellä sen sijaan tarkoitetaan virhetilanteiden oikeanlaista hallintaa.

- GET
 - Käytetään kun haetaan viestistä viestin tilaa ja resursseja
- PUT
 - Käytetään kun halutaan tehdä muutoksia resurssiin
- POST
 - Käytetään kun tehdään uusi resurssi tai kokoelma
- DELETE
 - Käytetään kun poistetaan kokoelma tai resurssi

Datapalvelun kannalta on tärkeää, että jokainen mainittu metodi toimii. Seurauksena käyttäjillä on mahdollisuus hakea, muuttaa, poistaa ja lisätä dataa. Käyttäjä ei kuitenkaan aina ole ihminen, vaan samat metodit tulee myös olla avoimet IoT-laitteille, mikäli niille on annettu siihen vaaditut oikeudet. API toimii niin ikään porttina käyttäjän (ihminen, kone tai jokin muu) ja datansäilöntäpaikan välillä.[56]

4.3.3. Verkko

Vesihuoltolaitokset suosivat LoRa-verkkoa, sillä verkko on todettu vesihuoltolaitosten kesken toimivaksi ja hyväksi ratkaisuksi. Ratkaisu ei myöskään rajoita myöhemmin uuteen verkkoteknologiaan siirtymistä [54]. Verkon valinnassa ei tällä hetkellä vesihuoltolaitosten kesken ole käytetty juurikaan muita verkkoratkaisuja kuin LoRa-teknologiaa. Datapalvelussa verkkoteknologia vaikuttaa lähinnä datan luotettavuuteen. Mikäli verkkoteknologia, jota käytetään, on stabiili ja luotettava, voidaan olettaa, että kerätty data on eheää.

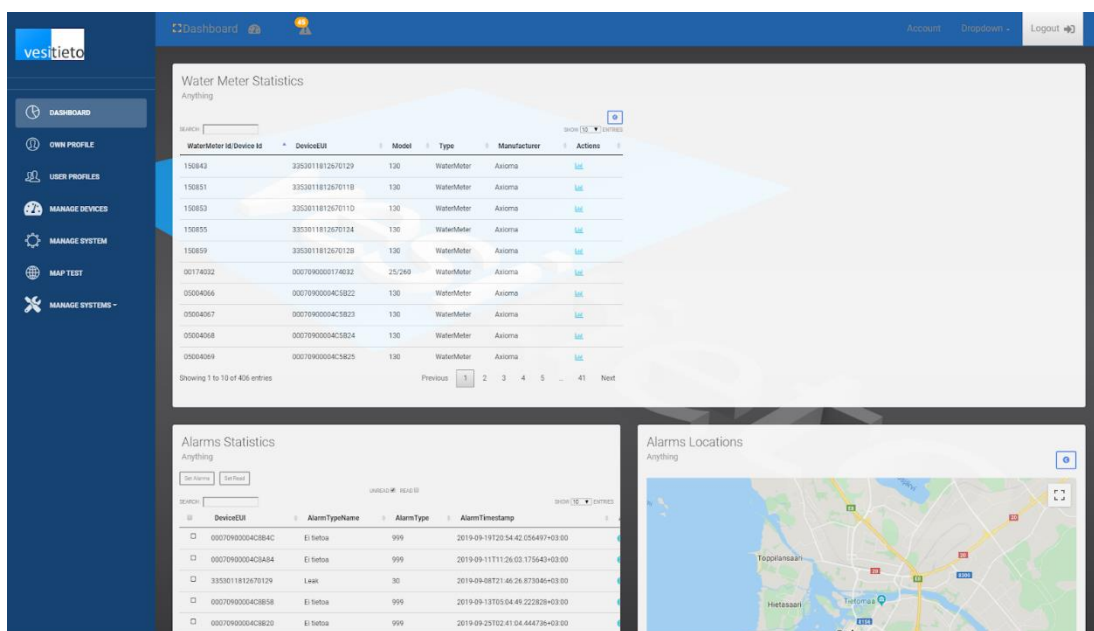
4.3.4. Kerätty data ja sen hyödyntäminen

IoT-laitteista kerätty data on niin kvalitatiivista kuin kvantitatiivista. Datapalvelu ottaa IoT-laitteesta datan vastaan ja tallentaa sen kokonaisuudessaan, mikäli laitteella on käyttöoikeudet Datapalveluun. Datasta hyödynnetään lähinnä kvantitatiivinen data, joka sijoitetaan edellä mainittuihin tietokantoihin. Kvantitatiivista dataa hyödyntämällä Datapalvelu näyttää asiakkaalle tietoja kerätystä datasta hyödyntäen graafista esitystä. Kvalitatiivisen datan hyödyntäminen on tällä hetkellä heikkoa, mutta parannuksia tulevaisuudessa on odotettavissa. Kvantitatiivista dataa hyödyntämällä Datapalvelussa voi nähdä kulutukseen liittyviä kaavioita ja taulukoita. Hälytykset on mahdollista nähdä taulukossa. Analyysiä varten käyttäjälle on annettu mahdollisuus järjestää dataa jokaisen taulukon sarakkeen mukaan (aikajärjestys, hälytystyyppi, yms.). Tällä tavoin Datapalvelu ei rajoita liiaksi asiakasta ja antaa toisaalta vapaat kädet tiedon analysointiin eri tarpeisiin.

4.4. Käyttöliittymä

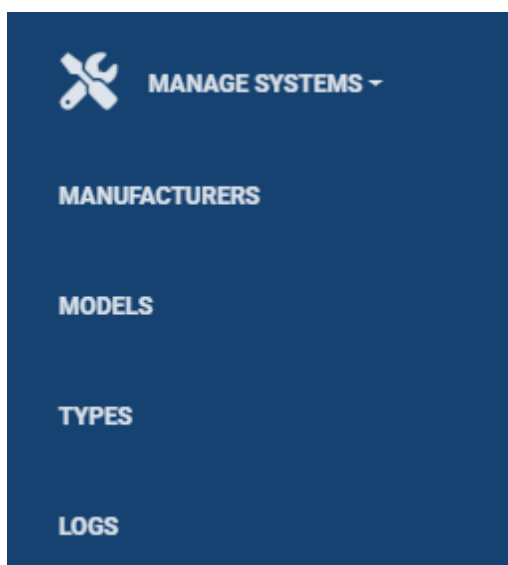
Käyttöliittymässä on käytetty neutraaleja värejä. Vesitiedon eri järjestelmissä sininen väri on dominoiva ja tämän käyttöä on jatkettu Datapalvelussa. Sininen väri ja sen

yleinen käyttö saa juurensa veden väristä (kuva 21). Käyttöliittymä on yksinkertainen ja selkeä. Samankaltaisia asioita tehdessä toiminnot ovat samantyyppisiä ja näin ollen käyttöönotto on helppoa ja käyttöliittymä on nopea oppia.



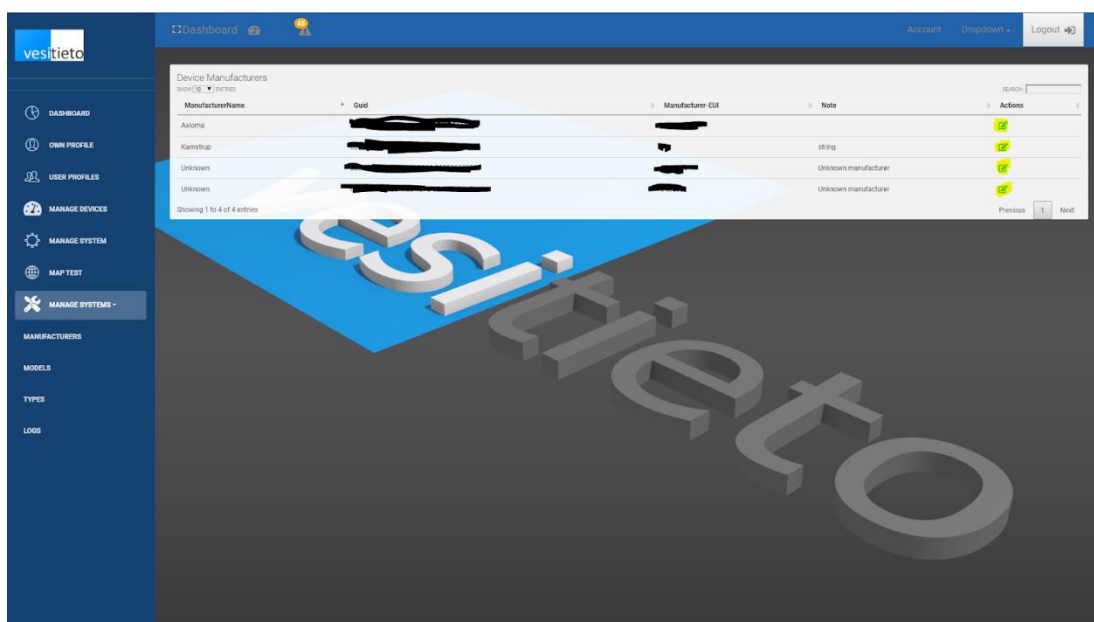
Kuva 21. Datapalvelu käyttöliittymä.

Käyttöliittymässä yläreuna sisältää toiminnot, jotka voidaan suorittaa nopeasti riippumatta siitä, millä välilehdellä ollaan. Esimerkiksi uloskirjautuminen on sijoitettu juuri oikeaan yläkulmaan, jotta uloskirjautuminen onnistuu milloin tahansa nopeasti. Eri järjestelmissä uloskirjautuminen on sijoitettu usein oikeaan yläkulmaan ja samaa sijoittelua käytetään Datapalvelussa. Näin vältetään ylimääräisiltä väärinymmärryksiltä ja palvelun käyttö on sujuvaa ja helppoa. Vasemmalle on sijoitettu toimintoja, joissa käyttäjä voi tehdä muutoksia ja seurata tarkemmin tiettyjä asioita. Esimerkiksi painamalla kuvan 21 sivupalkissa painiketta “MANAGE SYSTEMS” käyttäjälle avautuu oikeuksista riippuen kuvan 22 luettelo.



Kuva 22. Leikkaus käyttöliittymästä, hallitse järjestelmiä (Manage Systems).

Käyttäjälle on luotu mahdollisuus muokata itse muun muassa laitevalmistajien tietoja. Laitevalmistajat esitetään käyttäjälle listamuodossa, katso kuva 23.



Kuva 23. Leikkaus käyttöliittymästä, laitevalmistajat (Manufacturers).

Painamalla taulukossa oikealla olevaa kuvaketta (keltaisella maalattu), käyttäjä pääsee muuttamaan laitevalmistajalla olevia tietoja.

Käyttöliittymätyyppiratkaisua on käytetty kaikissa tiedoissa, joissa käyttäjälle on tarjottu mahdollisuus muokata järjestelmässä olevaa tietoa. Ensin valitaan suurempi kokonaisuus (tässä tapauksessa “manage systems”). Sen jälkeen alakokonaisuus (manufacturers, models, types, logs), josta päästään seuraamaan listamuodossa kaikkea alakokonaisuuden sisältämää tietoa. Taulukossa oikealla olevaa kuvaketta painamalla käyttäjä voi muokata haluttua tietoa.

4.4.1. Visualisointi ja työkalut

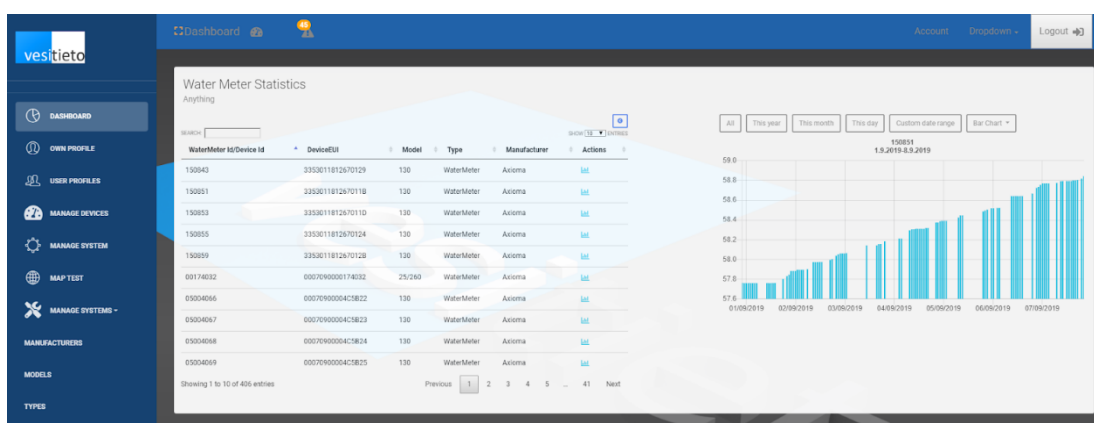
Visualisoinnissa on käytetty kaavioita, jotka ovat yleisesti tiedossa ja helposti ymmärrettävissä. Näihin lukeutuu kaksi yleisintä kaaviotyyppiä, joita ovat viivakaavio ja palkkikaavio. Oulun Vedellä kaavioita, joita käytetään muissa järjestelmissä, ovat viivakaavio, palkkikaavio, pylväskaavio ja piirakkakaavio. Kaikki neljä mainittua kaaviota tullaan toteuttamaan Datapalvelussa, kunhan resurssit, datan määrä ja aika antavat siihen mahdollisuuden. Suurena vaikuttavana tekijänä kaavioiden harvaan käyttöön löytyy datan määrästä. Datan määrä on kohtalaisen pientä vielä tässä vaiheessa projektia. Seurauksena ei kaikille kaaviotyypeille tässä vaiheessa Datapalvelun kehitystä löydy käyttötarkoitusta. Moniin eri kaaviotyypeihin on tutustuttu työssä, jotta tulevaisuudessa voidaan tarjota mahdollisimman monipuolinen ja kaiken kattava palvelu asiakkaille. Bullet-kaavio, laatikkokaavio, tutkakaavio ja vesiputouskaavio ovat ajankohtaisempia kaaviotyyppiä datamäärän kasvaessa ja asiakkaiden tarpeiden monipuolistuessa. Bullet-kaavio sopii, kun halutaan seurata tavoitteiden täyttymistä esimerkiksi liikevaihdon osalta. Laatikkokaavion käyttöä voidaan yhdistää vesimittarilukemien tiedusteluun sekä poikkeamien ja virheiden

nopeampaan löytämiseen. Vesiputouskaavion käyttö on ajankohtaisempaa, kun järjestelmään saadaan liitettyä laskutukseen liittyviä asioita, jolloin voidaan kuvata liikevaihtoa eri osa-alueilla. Tutkakaavio tuo omat hyvät puolensa esille, kun asiakkaalle halutaan kuvata montaa asiaa samassa kaaviossa ja tulkinta halutaan tehdä suuresta kokonaisuudesta vain yhden kaavion perusteella.

Visualisointiin on käytetty Chart.js-kirjastoa. Chart.js on Vesitiedolle tuttu kirjasto Vesitiedon tarjoamasta Vesitili-palvelusta. Muihin kirjastoihin (D3.js ja Dygraphs) on tutustuttu ja tehty taustatutkimusta tulevaisuutta ajatellen. Asiakkaiden tarpeiden kasvaessa useiden eri kirjastojen käyttö on ajankohtaisempaa, jotta voidaan tarjota monipuolisempia palveluita. Chart.js-kirjaston käyttöön on päädytty aikaisemman kokemuksen perusteella nopeuttamaan palvelun käyttöönottoa ja suunnitteluvaihetta.

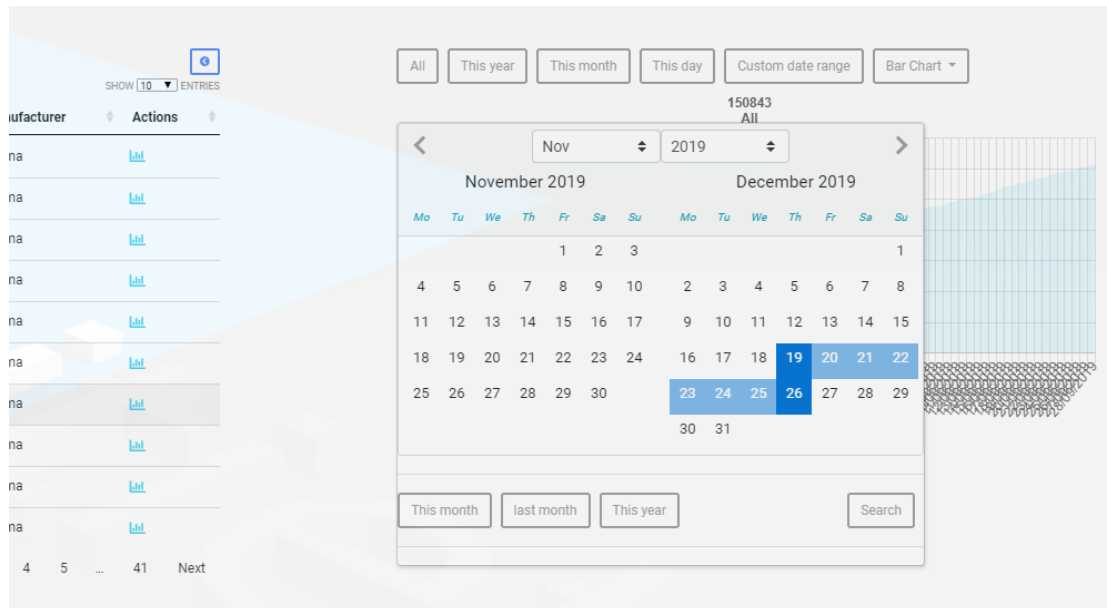
4.4.2. Kulutuksen seuranta

Kulutuksen seuranta on mahdollista mittarikohtaisesti. Kuvassa 24 taulukosta valitaan tietty mittari tarkasteluun. Mittarin valitseminen tapahtuu painamalla mittaria oikean rivin kohdalta. Mittarin valinnan jälkeen järjestelmä piirtää kulutuksesta palkkikaavion Chart.js-ohjelmakirjastoa hyödyntäen.



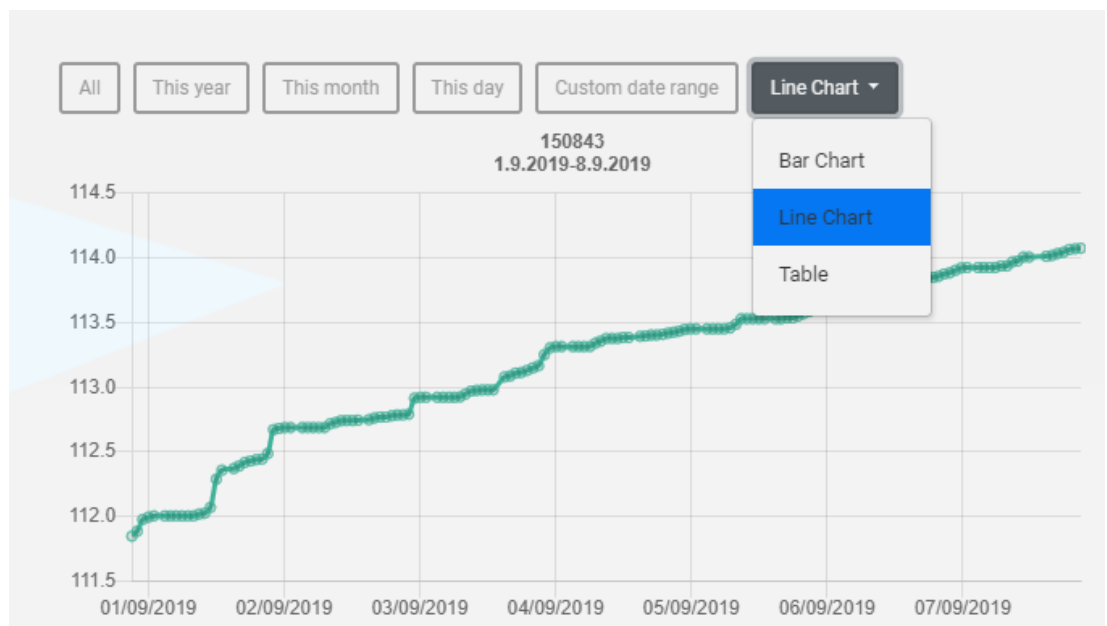
Kuva 24. Leikkaus käyttöliittymästä, kulutuksen seuranta.

Kun hiiri viedään palkin päälle, pieni laatikko näyttää kulutuksen tarkemmin. Kaavion yläreunassa näkyy tarkasteltu aikaväli ja mittarin numerosarja. Käyttäjä voi hakea kulutuksen haluamansa aikavälin mukaan tai käyttämällä valmiiksi annettuja aikavälejä, joita ovat kaikki, tämä vuosi, tämä kuukausi tai tämä päivä. Käyttäjän on siis mahdollista valita tarkasteluun myös tarkka päivämääräväli, kuten kuvassa 25.



Kuva 25. Leikkaus käyttöliittymästä, aikavälin asettaminen kulutuksen seurantaan.

Tarkempaan ja yksityiskohtaisempaa tarkastelua varten voidaan valita palkkikaavio, taulukkoratkaisu tai viivakaavio (kuva 26). Kun käyttäjä valitsee listasta uuden mittarin tarkasteluun, vanhat hakukriteerit säilyvät ja uuden mittarin tarkastelu tapahtuu samassa kaaviossa ja samalta aikaväliltä kuin edellisen mittarin tarkastelu. Toiminto helpottaa kulutuksen seuranta samoilla hakukriteereillä.

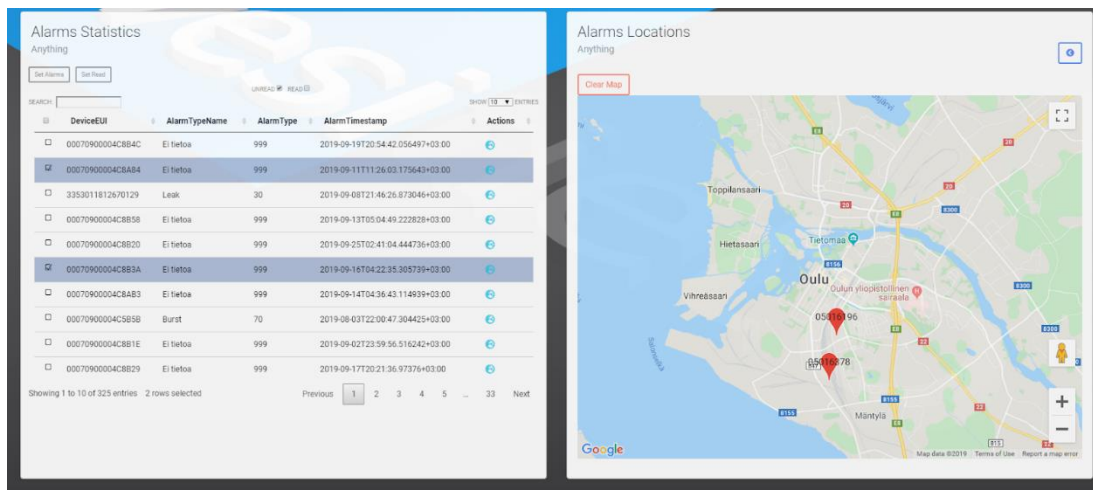


Kuva 26. Leikkaus käyttöliittymästä, kulutuksen seuranta ja kaaviot.

Kulutuksen seurannasta, niissä käytetyistä kaavioista ja toiminnoista on tullut ainoastaan positiivista palautetta. Palaute on peräisin paitsi eri vesihuoltolaitosten henkilöiltä, mutta myös Onniselta sekä Digitalta.[53]

4.4.3. Hälytykset

Laitteiden lähetettyjä hälytyksiä voidaan tarkastella taulukossa sekä kartalla. Taulukkoon tuodaan hälytykset, joita ei ole merkitty luetuiksi, kuva 27. Vaihtamalla asetuksia on käyttäjän mahdollista tarkastella myös luettuja hälytyksiä.

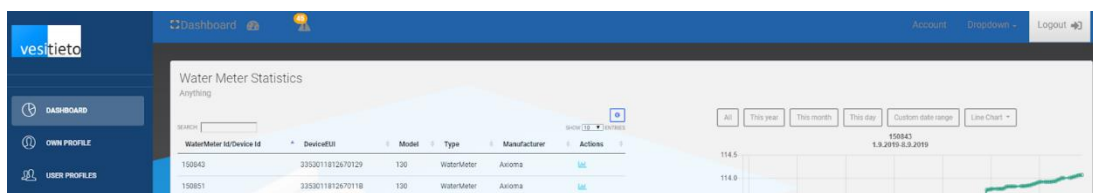


Kuva 27. Leikkaus käyttöliittymästä, hälytykset.

Valitsemalla hälytys ja painamalla kuittaus-nappia vasemmasta yläkulmasta, hälytys kuittaantuu nähdynä. Hälytyksille on lisätty karttanäkymä, josta nähdään hälytyksen sijaintitiedot. Painamalla maapallon kuvaa taulukosta, asettaa järjestelmä kartalle hälytyksen sijaintitiedot. Painamalla merkkiä, jonka järjestelmä tuo kartalle, käyttäjä voi tarkastella lisätietoja hälytyksestä. Set Alarm-painike asettaa kartalle kaikki luetut, lukemattomat tai riippuen asetuksista kaikki hälytykset kartalle. Kuvasta 28 nähdään, ettei hälytysten priorisointiin (Alarm Type) käytetä värejä vaan ainoastaan numerokoodia. Hälytykset on mahdollista järjestää järjestykseen kaikkien sarakkeiden mukaan painamalla sarakkeesta. Käyttäjälle tarjotaan myös "search" toimintoa, joka suodattaa tulokset. Kirjoittamalla "search" kenttään 999, näyttää taulukko rivit, joista löytyy arvo 999 jostakin sarakkeesta.

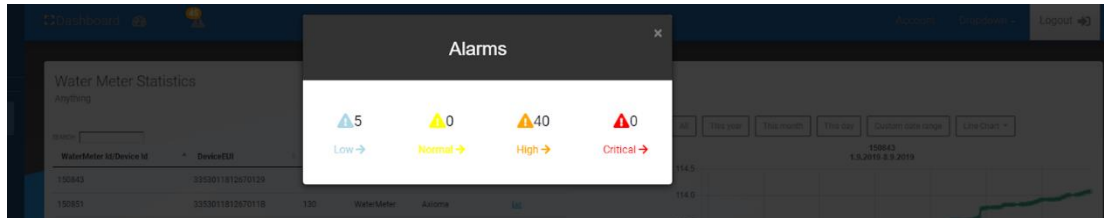
Palaute hälytysten seurannasta on ollut positiivista, mutta parannuksia on toivottu. Erityisesti hälytysten erottelu kriittisten ja ei niin kriittisten hälytysten välillä halutaan selvemmäksi. Oulun Vedellä kartan hyödyntämistä osana palvelua kehuttiin erittäin hyväksi ominaisuudeksi. Kartan hyödyntämiseen tullaan keskittymään lisää tulevaisuudessa, sillä mielenkiinto kartan hyödyntämiselle asiakkaiden parissa on ollut suurta, etenkin Oulun Vedeltä.

Hälytysten priorisointia väreillä on aloitettu (kuva 28). Yläreunassa kolmio, jossa näkyy oranssi väri ja numero 45, kuvaa hälytysten määrää. Kolmio muuttaa väriään riippuen lukemattomista hälytystyypeistä.



Kuva 28. Leikkaus käyttöliittymästä, hälytys ikoni.

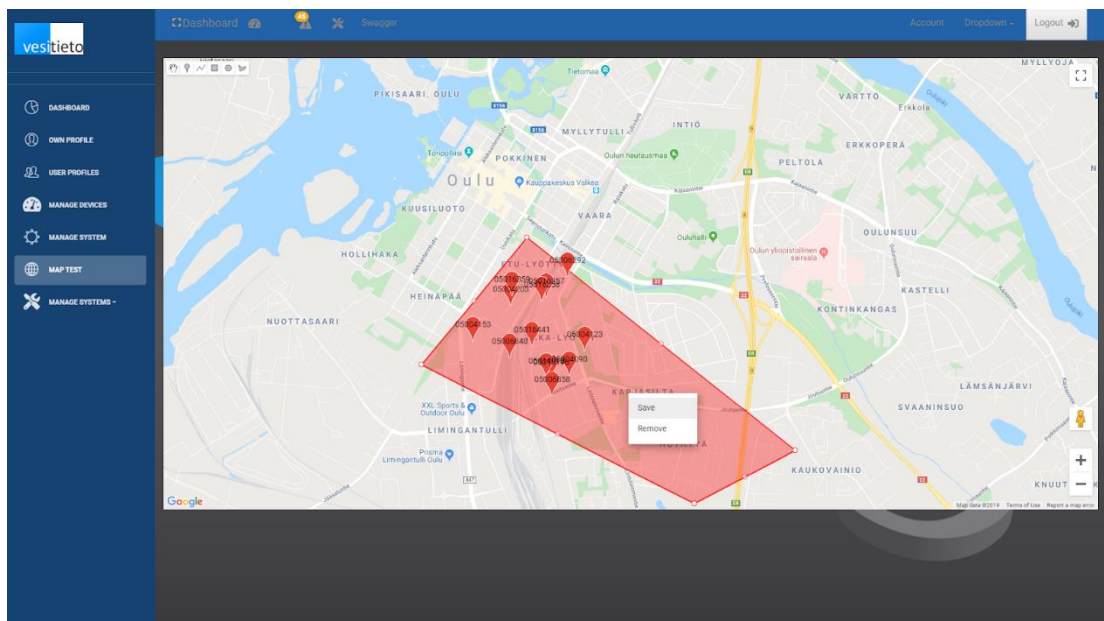
Painamalla kolmioikonia aukeaa ponnahdusikkuna, jossa on eroteltuna hälytykset tyypeittäin ja hälytysten määrät (Kuva 29). Tausta on himmennetty, jotta käyttäjä voi keskittyä avattuun asiaan tarkemmin. Käyttäjä siirtyy suoraan hälytyksiin, valitsemalla jonkin hälytystyyppin suodatuskriteeriksi. Tausta muuttuu normaaliksi ja himmennys poistuu, kun käyttäjä sulkee ikkunan tai siirtyy seuraamaan hälytyksiä painamalla hälytystyyppiä.



Kuva 29. Leikkaus käyttöliittymästä, hälytykset eroteltu.

4.4.4. Alueiden määrittely piirtämällä

Alueiden piirtäminen ja määrittely kartalle on kehitysvaiheessa. Toiminnalle on nähty suurta potentiaalia ja sen seurauksena toimintaa on lähdetty kehittämään. Kartalle piirretään alue, jonka jälkeen järjestelmä etsii automaattisesti kaikki alueella sijaitsevat laitteet (Kuva 30). Hiiren oikeaa näppäintä painamalla käyttäjä voi tallentaa tai poistaa piirretyn alueen.



Kuva 30. Leikkaus käyttöliittymästä, alueiden määrittely ja tallennus.

Toiminto on keskeneräinen, mutta toteutetaan ennen tuotantoon siirtymistä. Vesitieto tarvitsee lisää informaatiota siitä, mitä alueellista tietoa halutaan näyttää. Toimintoa tullaan kehittämään lisää erityisesti asiakkaiden toiveiden mukaisesti.

4.5. Tulevaisuus

Potentiaalisten asiakkaiden toiveiden tiedustelua jatketaan ja toimintoja kehitetään palautteen mukaan. Toteutus vaatii yhä tiiviimpää yhteistyötä asiakkaiden kanssa sekä tarkempaa tarpeiden tiedustelua. Ensisijaisesti lähdetään kehittämään kartan hyödyntämistä, sisältäen alueiden piirtämisen ja määrittelyn karttatyökaluja hyväksi käyttäen. Kartan rinnalla kehitetään hälytysten seuraamista ja priorisointia eteenpäin. Suunnitelmiin kuuluu monipuolisempi kaavioiden käyttö ja laajemman kokonaisuuden tarkastelu sekä tiedustelu. Kaavioiden käyttö on riippuvainen asiakkaiden toiveista ja tällä hetkellä suosituimmat kaaviot ovat palkkikaavio, piirakkakaavio sekä viivakaavio.

Piirakkakaaviolle todennäköisin käyttötarkoitus löytyy, kun lähdetään seuraamaan alueellista kulutusta. Alueellinen kulutus näytettäisiin tällöin piirakkakaavion muodossa. Kulutuksen vertailuun kuukausitasolla on pohdittu viiva- ja palkkikaavion yhdistelmää. Palkkikaavio kuvaa kulutusta tällä hetkellä ja viivakaavio viime vuoden kulutusta samalta aikaväliltä. Ratkaisu vaatii suuremman aikavälin dataa, joten toiminto on ajankohtainen myöhemmässä vaiheessa. Tällöin puhutaankin todennäköisesti massadatasta.

Tulevaisuuden suunnitelmiin kuuluu myös vesihuollon lisäksi muihin energian aloihin laajentuminen. Datapalvelun toivotaan tarjoavan kustannustehokas ja yritystoimintaa tehostava järjestelmä myös muille energia-aloille. Järjestelmän tuotantoversio on suunnattu helmikuulle 2020. Aikataulu on suuntaa antava ja mahdollisia myöhästymisiä saattaa aiheutua. Datapalvelua tullaan kehittämään päivittäin myynnin ohessa. Uudet toiminnot ja ominaisuudet kehittyvät asiakkaiden tarpeiden kasvaessa.

5. TULOSTEN TARKASTELU

Työssä tutkitaan vesihuollon ongelmia ja mahdollisuuksia, joita digitalisoituminen tuo mukanaan. Ongelmien pohjalta on kehitetty täysin uusi järjestelmä Datapalvelu, joka helpottaa digitalisoitumista. Datapalvelu tarjoaa vesihuollon asiakkaille palvelun, jota käyttämällä uusien etäluettavien mittarien datan tallentaminen sekä hallitseminen on helppoa, turvallista ja nopeaa. Datapalvelu tarjoaa asiakkaalle mahdollisuuden tarkastella kerättyä dataa ja siitä valmiiksi tehtyä analysointia muun muassa kaavioiden muodossa. Ratkaisuisissa on pyritty muistamaan, että myöhemmin käsitellään massadataa.

Järjestelmälle on tarvetta, sillä datan tallentamiseen ja mallintamiseen ei löydy järjestelmää, joka olisi suunniteltu etäluettaville mittareille, saati massadatalle. Ongelmiin lähdettiin hakemaan ratkaisuja Datapalvelun kautta. Datapalvelu on kehitysvaiheessa ja ensimmäisiä versioita pyritään esittämään mahdollisille asiakkaille lähitulevaisuudessa. Tällä tavoin saadaan suoraa palautetta järjestelmän toiminnoista sekä ulkoasusta. Palaute on ollut pääosin positiivista ja järjestelmälle nähdään asiakkaiden keskuudessa paljon potentiaalia. Tarve järjestelmälle kasvaa lisää, kun vesihuolto siirtyy yhä enemmän kohti etäluettavia mittareita ja kun data määrät kasvavat. Datapalvelun kehitys on onnistunut hyvää vauhtia ja ensimmäisiä tuotantoon otettavia versioita voidaan olettaa vuoden 2020 alkupuolella.

Työ antaa laajan kuvan Datapalvelussa olevista ominaisuuksista ja siinä käytetyistä visuaalisista menetelmistä. Vahvuuksiin kuuluu erityisesti tiivis yhteistyö vesihuoltolaitosten kanssa ja tehokas ongelmanratkaisukyky. Tällä tavoin pystytään vastaamaan uusiin ilmestyviin tarpeisiin nopeasti ja oikeanlaisilla ratkaisuilla. Heikkouksiin kuuluu datamäärän pienuus. Etukäteen on erittäin vaikea tehdä ratkaisuja pohjautuen suureen datamäärään, koska sitä ei ole saatavilla. Tähän on varauduttu hyvin käyttämällä ohjelmakirjastoja ja työkaluja, jotka tukevat paitsi suuria määriä dataa, mutta erityisesti massadataa. Muun muassa visualisoinnissa on käytetty kaavioita, joiden tiedetään tukevan massadataa ja skaalautuvan hyvin. Tarkastellessa tavoitteita (UI kehitys ja suunnittelu, datan analysointi, datan visualisointi, datan jalostus) tämän työn osilta onnistuttiin hyvin niin suunnittelussa kuin kehityksessä.

- UI kehitys ja suunnittelu
 - UI on selkeä ja helppokäyttöinen. Asiakkaat ovat antaneet positiivista palautetta sekä selkeydestä että toiminnoista, jotka näyttävät nykyaikaisilta. Nopea ja suoraviivainen käyttö on saatu aikaan käyttämällä samoja toimintaperiaatteita joka paikassa
- Datan analysointi
 - Datan analysoinnissa on päästy alkuun. Analyysia näkyy kaavioissa, mutta käyttäjälle annetaan mahdollisuus tehdä analyysia kaavioiden ja kartan perusteella myös itse. Käyttäjälle on annettu mahdollisuus järjestää data eri kriteerien mukaan ja suodattaa osa datasta hetkellisesti.
- Datan visualisointi
 - Visualisointiin on tutkittu rutkasti eri vaihtoehtoja, joita voidaan käyttää datan määrän kasvaessa ja syvällisempien analyysitarpeiden kasvaessa. Yleisistä kaaviotyypeistä käytetään palkki- ja viivakaavioita sekä lisäksi hyödynnetään karttaa ja sen tuomia ominaisuuksia.

- Datan jalostus
 - Dataa on voitu muuttaa tiedoksi muun muassa vedenkulutuksen laskentaa varten. Vesihuollon tarpeiden kasvaessa datan jalostusta kehitetään eteenpäin, kun tiedetään tarkemmin, mitä tietoa datasta halutaan ja minkälaisessa muodossa.

Datapalvelu tarjoaa vesihuollon digitalisoitumiseen ratkaisuja ja sen kehitys jatkuu edelleen. Vesihuollon tarpeisiin vastataan tällä hetkellä erittäin hyvin, ja lähitulevaisuudessa Datapalvelu tarjoaa ratkaisun IoT-datan tallentamiseen, analysoimiseen, havainnollistamiseen ja visualisoimiseen. Työn oli tarkoitus olla pohja Datapalvelun suunnittelulle ja rakentamiselle hyödyntäen visuaalisia ratkaisuja, kuten kaavioita.

6. YHTEENVETO

Vesihuolto on paljon jäljessä digitalisoitumisessa verrattuna muihin energia-aloihin. Vesihuollolla ei tällä hetkellä ole tarkkaa tietämystä siitä, kuinka nopeasti he tulevat uusimaan vanhat mittarit etäluettaviin mittareihin. Vesihuollon parissa on vielä monta kysymysmerkkiä, mutta halu digitalisoitua on suuri ja toteutuu erittäin todennäköisesti lähivuosina. Mikäli Vesitiedon kehittämä Datapalvelu valitaan osaksi vesihuollon järjestelmää, Datapalvelu voi toimia datan tallennus- ja hallitsemispaikkana sekä tuottaa tallennetusta datasta analyysia asiakkaille. Tarjotun informaation avulla vesihuollon on mahdollista seurata ja arvioida tulevaisuutta paremmin ja tarkemmin sekä tehdä kriittisiä päätöksiä entistä tehokkaammin, nopeammin ja turvallisemmin.

Tämä työ käsittelee IoT-mittareihin suunniteltua ja kehitettyä Datapalvelua. Työssä käydään läpi, kuinka Datapalvelu on suunniteltu käyttäjälle ja mitä asioita on otettu huomioon erityisesti toiminnoissa, ulkoasussa ja graafisissa esityksissä. Työ käy läpi menetelmät, joita on käytetty Datapalvelun rakentamiseen ja käyttäjäystävällisen käyttöliittymän toteuttamiseen. Datan analysointia varten käyttäjälle on annettu useita eri vaihtoehtoja. Kaaviot antavat nopeasti trendejä eri asioista kuten kulutuksesta. Kulutusta on helppo seurata ja trendissä näkyvät poikkeamat on helppo huomata. Käytetyt kaaviotyypit ovat erittäin yleisesti käytettyjä ja näin ollen lähes jokaisen käyttäjän helposti tulkittavissa. Kaaviot ovat selkeitä, helppolukuisia ja helposti tulkittavissa.

Palkki- ja viivakaavion lisäksi palvelua on täydennetty kartan käyttämisellä, jonka avulla voidaan tarkastella tarkemmin mallinnettuja alueita ja mittareiden sijaintia. Tulevaisuudessa käyttäjä voi myös tarkastella määritettyjä alueita kartalta ja niiden sisältämää tietoa, kuten määritetyn alueen vedenkulutusta. Datapalvelu analysoi myös IoT-mittareiden tuottamia hälytyksiä. Hälytyksistä annetaan käyttäjälle tietoa esimerkiksi siitä, mikä on hälytystyyppi, hälytyksen numero, mittarinnumero. Lisäksi on mahdollisuus paikantaa mittari kartalla, mikäli mittarille on annettu koordinaatit. Hälytykset on mahdollista asettaa luetuksi, jolloin ne eivät näy listassa. Käyttäjäystävällisyyden takaamiseksi myös jo luetut hälytykset saa takaisin listaan, jolloin voidaan tarkastella mittalaitteiden hälytyshistoriaa pitemmällä aikavälillä. Hälytykset, siinä missä kaiken muunkin, voi järjestää eri kriteerien mukaan kuten aikajärjestykseen tai järjestykseen hälytystyyppin mukaan. Näin käyttäjällä on mahdollisuus tarkastella hälytyksiä juuri haluamallaan tavalla.

Kun tarkastellaan Datapalvelua tämän työn osalta, Datapalvelu tarjoaa kattavan ohjelmiston datan seurantaan ja analysointiin samalla hyödyntäen visuaalisia ratkaisuja. Tarkastelua ja analysointia voidaan suorittaa monin eri keinoin, kuten kaavioita seuraamalla. Tuotetun informaation avulla voidaan helpottaa ja nopeuttaa päätöksentekoa vesihuollon parissa.

7. LÄHTEET

- [1] Allied ICT Finland (AIF) Water Ecosystem (2019) Työpaja Oulun Veden ja yritysten kanssa. URL: <https://www.water-solutions.org/>.
- [2] Ojala, M. (2019) Toimitusjohtaja Markku Ojalan haastattelu 2019. Vesitieto Oy, Oulu.
- [3] Ashton, K. (2009) That ‘internet of things’ thing. RFID journal, 22(7), 97-114.
- [4] Madakam, S. (2015) Internet of things: smart things. International journal of future computer and communication, 4(4), 250.
- [5] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013) Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems, 29(7), 1645-1660. DOI: <https://doi.org/10.1016/j.future.2013.01.010>.
- [6] Evans, P. C., & Annunziata, M. (2012) Industrial internet: Pushing the boundaries. General Electric Reports, 488-508.
- [7] Li, P., Li, J., Huang, Z., Gao, C. Z., Chen, W. B., & Chen, K. (2018) Privacy-preserving outsourced classification in cloud computing. Cluster Computing, 21(1), 277-286. DOI: <https://doi.org/10.1007/s10586-017-0849-9>.
- [8] Saturno, M., Pertel, V. M., Deschamps, F., & Loures, E. D. F. (2017) Proposal of an automation solutions architecture for industry 4.0. In Proceedings of the 24th International Conference on Production Research. Poznan: ICPR.
- [9] Boyes, H., Hallaq, B., Cunningham, J., & Watson, T. (2018) The industrial internet of things (IIoT): An analysis framework. Computers in Industry, 101, 1-12. DOI: <https://doi.org/10.1016/j.compind.2018.04.015>.
- [10] Marr, B. (2016) What everyone must know about industry 4.0. URL: <https://www.forbes.com/sites/bernardmarr/2016/06/20/what-everyone-must-know-about-industry-4-0/6d63e914795f>, Viitattu 20.10.2019.
- [11] Augustin, A., Yi, J., Clausen, T., & Townsley, W. (2016) A study of LoRa: Long range & low power networks for the internet of things. Sensors, 16(9), 1466. DOI: <https://doi.org/10.3390/s16091466>.
- [12] Holtzblatt, K., & Beyer, H. (2016) Contextual design: Design for life. Morgan Kaufmann.
- [13] Silverman, D. (2013) Doing qualitative research: A practical handbook. SAGE publications limited.

- [14] Green, J., & Thorogood, N. (2018) *Qualitative methods for health research*. Sage.
- [15] Hammersley, M., & Atkinson, P. (2007) *Ethnography: Principles in practice*. Routledge.
- [16] Coghlan, D. (2019) *Doing action research in your own organization*. SAGE Publications Limited.
- [17] Britten, N. (2006) Qualitative research in health care. *Qualitative research in health care*, 12-20.
- [18] Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008) Methods of data collection in qualitative research: interviews and focus groups. *British dental journal*, 204(6), 291.
- [19] Kvale, S., & Brinkmann, S. (2009) *Interviews: Learning the craft of qualitative research interviewing*. Sage.
- [20] Stewart, D. W., & Shamdasani, P. N. (2014) *Focus groups: Theory and practice* (Vol. 20). Sage publications.
- [21] Bloor, M. (Ed.). (2001) *Focus groups in social research*. Sage.
- [22] Queirós, A., Faria, D., & Almeida, F. (2017) Strengths and limitations of qualitative and quantitative research methods. *European Journal of Education Studies*. DOI: <https://doi.org/10.5281/zenodo.887089>.
- [23] Hammersley, M. (2016) *Reading ethnographic research*. Routledge.
- [24] Patton, M. Q., & Patton, M. Q. (2002) *Qualitative research and evaluation methods*. Thousand Oaks, Calif: Sage Publications.
- [25] Treiman, D. J. (2014) *Quantitative data analysis: Doing social research to test ideas*. John Wiley & Sons.
- [26] Senders, J. W., & Moray, N. P. (1991) *Human Error: Cause. Prediction, and Reduction*, Lawrence Erlbaum: Hillsdale, NH.
- [27] Measurement Error. (2016) ULR: https://serc.carleton.edu/quantskills/teaching_methods/und_uncertainty/measurement_error.html. Viitattu 12.11.2019.
- [28] Rowley, J. (2007) The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180. DOI: <https://doi.org/10.1177/0165551506070706>.
- [29] Ware, C. (2012) *Information visualization: perception for design*. Elsevier.

- [30] The Data Visualization Handbook. (2019). URL: <https://www.sap.com/documents/2018/04/8cb8821e-fc7c-0010-87a3-c30de2ffd8ff.html>. Viitattu 15.10.2019.
- [31] Keim, D., Qu, H., & Ma, K. L. (2013) Big-data visualization. IEEE Computer Graphics and Applications, 33(4), 20-21.
- [32] Koponen, J., & Hildén, J. (2019) Data visualization handbook. Aalto korkeakoulusäätiö.
- [33] Ware, C. (2010) Visual thinking: For design. Elsevier.
- [34] Nayak, A. (2019) Introduction to Data Visualization. URL: <https://www.udemy.com/introduction-to-data-visualization/>. Viitattu 20.11.2019.
- [35] Wurman, R. S. (1989) Information anxiety. NY: Doubleday.
- [36] Kuusela, V. (2000) Tilastografiikan perusteet. Edita.
- [37] Porter, M. M., & Niksiar, P. (2018) Multidimensional mechanics: Performance mapping of natural biological systems using permuted radar charts. PloS one, 13(9). DOI: <https://doi.org/10.1371/journal.pone.0204309>.
- [38] Häiriöt kartalla. Oulun Vesi. URL: <https://kartta.ouka.fi/keskeytyskartta/?language=fin>, Viitattu 30.09.2019.
- [39] Wikipedia. Wikipedia Commons. URL: <https://commons.wikimedia.org/wiki/File:Bullet-bar-terms-key.png>. Viitattu 09.10.2019.
- [40] Gillespie, T. W. (2012) Understanding waterfall plots. Journal of the Advanced Practitioner in Oncology, 3(2), 106.
- [41] Box and Whisker Plot. (2019). URL: https://datavizcatalogue.com/methods/box_plot.html. Viitattu 11.10.2019.
- [42] Chart.js. (2019) URL: <https://www.chartjs.org/>. Viitattu 13.11.2019.
- [43] De Rosa, A. (2015) Creating Beautiful Charts with Chart.js. URL: <https://www.sitepoint.com/creating-beautiful-charts-chart-js/>. Viitattu 15.11.2019.
- [44] Bostock, M. (2019) Data-Driven Documents. URL: <https://d3js.org/> Viitattu 13.11.2019.
- [45] dygraphs. (2019) URL: <http://dygraphs.com/>. Viitattu 15.11.2019.

- [46] Kamstrup. (2019) Älymittausratkaisut.
URL: https://www.kamstrup.com/fi-fi?gclid=Cj0KCQiAuefvBRDXARIsAFEOQ9HezfnDa0uYWZIl1sD_sTMkOFRZwT-flWN-AQFC0O-lcfKben_uHV5AaAsw7EALw_wcB&gclidsrc=aw.ds. Viitattu 8.10.2019.
- [47] Salo, I. (2014) Big data & pilvipalvelut. Jyväskylä: Docendo.
- [48] Jokio J (2020) Back-end Reference Architecture for Smart Water Meter Data Gathering Service. Master's thesis. University of Oulu, Department of Computer Science and Engineering.
- [49] Snijders, C., Matzat, U., & Reips, U. D. (2012) " Big Data": big gaps of knowledge in the field of internet science. International Journal of Internet Science, 7(1), 1-5.
- [50] Kaplan, A., & Haenlein, M. (2019) Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons, 62(1), 15-25. DOI: <https://doi.org/10.1016/j.bushor.2018.08.004>.
- [51] Kepner, J., Gadepally, V., Michaleas, P., Schear, N., Varia, M., Yerukhimovich, A., & Cunningham, R. K. (2014, September) Computing on masked data: a high performance method for improving big data veracity. In 2014 IEEE High Performance Extreme Computing Conference (HPEC) (pp. 1-6). IEEE.
DOI: <https://doi.org/10.1109/HPEC.2014.7040946>.
- [52] FINLEX ® - Ajantasainen lainsäädäntö: Vesihuoltolaki 119/2001 (2001)
URL: [https://www.finlex.fi/fi/laki/ajantasa/2001/20010119?search\[type\]=pika&search\[pika\]=vesihuolto](https://www.finlex.fi/fi/laki/ajantasa/2001/20010119?search[type]=pika&search[pika]=vesihuolto). Viitattu 1.10.2019.
- [53] Kaijalainen, H. (2019) Myyntijohtaja Harri Kaijalaisen haastattelu 2019. Vesitieto Oy, Oulu.
- [54] Kilpeläinen, T. (2019) Verkostopäällikkö Tero Kilpeläisen haastattelu 4.12.2019. Oulun Vesi, Oulu.
- [55] Alanära, S. (2019) Kehittämispäällikkö Sara Alanärän haastattelu 4.12.2019. Oulun Vesi, Oulu.
- [56] Masse, M. (2011) REST API Design Rulebook: Designing Consistent RESTful Web Service Interfaces. " O'Reilly Media, Inc."